A Quasi-Hamiltonian Discretization of the Thermal Shallow Water Equations

Christopher Eldred[a,*], Thomas Dubos[b], Evaggelos Kritsikis[c]

[a] *Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France*
[b] *Laboratoire de Météorologie Dynamique/IPSL, École Polytechnique, 91120 Palaiseau, France*
[c] *Laboratoire d'Analyse, Géométrie et Applications, Université Paris 13, 93430 Villetaneuse, France*

Abstract

The rotating shallow water (RSW) equations are the usual testbed for the development of numerical methods for three-dimensional atmospheric and oceanic models. However, an arguably more useful set of equations are the thermal shallow water equations (TSW), which introduce an additional thermodynamic scalar but retain the single layer, two-dimensional structure of the RSW. As a stepping stone towards a three-dimensional atmospheric dynamical core, this work presents a quasi-Hamiltonian discretization of the thermal shallow water equations using compatible Galerkin methods, building on previous work done for the shallow water equations. Structure-preserving or quasi-Hamiltonian discretizations methods, that discretize the Hamiltonian structure of the equations of motion rather than the equations of motion themselves, have proven to be a powerful tool for the development of models with discrete conservation properties. By combining these ideas with an energy-conserving Poisson time integrator and a careful choice of Galerkin spaces, a large set of desirable properties can be achieved. In particular, for the first time total mass, buoyancy and energy are conserved to machine precision in the fully discrete model.

Keywords: thermal shallow water equations, dynamical core, mixed finite elements, finite element exterior calculus, mimetic Galerkin differences, Hamiltonian mechanics

1. Introduction

The rotating shallow-water (RSW) equations are a useful model in their own right for a diversity of natural flows. This has motivated many previous authors to propose and study numerical methods for the RSW. They are also of interest as an intermediate model towards three-dimensional modeling of atmospheric and oceanic circulations. A further step along the path towards the full equations is the thermal shallow water equations (also known as the Ripa equations) [1], which augment the shallow water equations with an additional field akin to temperature or entropy that is transported and modifies the dynamics through its effects on the fluid density and the horizontal pressure gradient. Indeed, when expressed in a floating

*Corresponding Author: Christopher Eldred (chris.eldred@gmail.com)

Lagrangian vertical coordinate, rather than the usual Eulerian vertical coordinate, three-dimensional equations of oceanic and atmospheric motion (see, e.g. [2]) closely resemble a system of multiple layers obeying TSW-like dynamics and coupled only by the pressure force. Additionally, the thermal shallow water equations and the fully compressible equations in both Eulerian and generalized vertical coordinates share similar Hamiltonian structures [3], with near-identical Poisson brackets and differing only in the Hamiltonian. Compared to the RSW equations, the TSW equations have attracted relatively little attention [4, 5], and few numerical solutions have ever been produced (although see [6, 7, 8]). This is especially true in the atmospheric and oceanic dynamical core development communities, particularly when considering structure-preserving discretizations. Here, we are interested in solving the TSW equations as a milestone towards solving three-dimensional equations and as a testbed to demonstrate novel ideas elaborating on recent work on structure-preserving discretization for the RSW equations [9, 10, 11].

Regarding the RSW equations, previous work on numerical methods fall roughly in two categories. In the first category, the hyperbolic structure of the equations expressed in flux-form is exploited, leading generally to schemes of finite-volume or discontinuous-Galerkin type. Typical issues to be dealt with are the design of low-dissipation Riemann solvers, the preservation of steady states in the presence of source terms due to bottom topography, or the stability of extensions to higher order. In the second category, to which the present work belongs, the curl-form of the equations (akin to the Crocco equation for compressible fluid) is exploited, leading generally to finite-difference [12, 13, 14, 15, 16, 17] or finite-element schemes, including spectral elements [18] and mixed finite elements [9, 10, 19, 20, 21, 22]. The latter have paved a way towards higher-order spatial accuracy, which for finite-difference methods is limited to first or second order at best, especially on irregular meshes [23]. Typical concerns (a more exhaustive list of desirable numerical properties is discussed in Section 3.1) are then the possibly anomalous numerical dispersion of waves [24, 25, 26, 27, 28, 29, 30, 31, 32], the preservation of steady states, especially geostrophic balance [33, 34], the numerical dynamics of potential vorticity [35, 36], the Hollingsworth instability [37, 38, 39, 40, 41] and the exact conservation of discrete energy and enstrophy [10, 12, 13, 14, 16, 17, 19, 42, 43, 44].

The desire to conserve exactly a discrete approximation of total energy stems from the ultimate goal to simulate atmospheric and oceanic flows over long time scales, an exercise which can be imperiled by the slow accumulation of tiny conservation errors. There is a long history of finite-difference schemes successfully exploiting the curl-form to achieve energy conservation up to time discretization errors. To date however, the majority of such successes have been singular achievements due to the intuition and ingenuity of the authors, without a clear path towards the systematic design of energy-conserving schemes. Probably, this is because it has taken time to realize that the RSW equations conserve total energy because they derive from a variational principle, which also implies the existence of the curl-form, itself reflecting a non-canonical Hamiltonian structure (see Section 2). Following this recognition, it has been suggested that this structure could be imitated at the discrete level to systematically design energy-conserving spatial discretizations, so-called quasi-Hamiltonian schemes [43]. Still, more than thirty years later, this strategy remains the

exception [14, 42, 45, 46, 47] rather than the rule. Apart from those exceptions, recent work on energy-conserving discretizations either ignores altogether the connection between the curl form and the Hamiltonian structure [19, 48] or identifies a quasi-Hamiltonian structure after the fact [10][1].

The present work defines, implements and evaluates structure-preserving finite element schemes for the TSW along the lines of recent work for the RSW [10] and innovates over it in four main areas :

1. Rather than the RSW, we address the TSW whose structure more closely resembles that of three-dimensional fluid equations and which have received less attention; especially, several possible choices for the discretization of the additional prognostic field of the TSW are considered theoretically and the consequences of these choices are discussed based on numerical experiments,

2. In addition to previously considered combinations of compatible mixed finite element spaces, which lead to certain dispersive anomalies when applied to the linearized RSW equations [26, 27, 31], we consider the recently proposed Mimetic Galerkin Difference elements [49], which have been shown to be free of those dispersive anomalies [24],

3. The Hamiltonian structure of the TSW equations is the starting point of our discretization procedure, leading systematically to energy-conserving quasi-Hamiltonian spatial discretizations, while still leaving some free choices whose numerical consequences are studied,

4. In a major step forward, we leverage the quasi-Hamiltonian structure of our spatial discretizations and recently proposed temporal discretization schemes to achieve exact conservation of discrete energy (along with mass and total buoyancy) for all variants of the discretization, while previous schemes conserve energy only up to time discretization errors; the resulting scheme, while implicit in time, is arguably not significantly more expensive than similar schemes currently used for numerical weather forecasting.

This work is presented as follows. In Section 2, we present the TSW and linearized TSW equations in various forms. We focus on the curl-form equations, which we reformulate in terms of the Poisson brackets that give them their Hamiltonian structure. In Section 3, Mimetic Galerkin Difference spaces are presented and a general procedure for finite-element spatial discretization is developed. The fundamental idea is to discretize the Poisson brackets and the total energy (Hamiltonian) themselves. This systematic procedure yields weak forms of the equations with a quasi-Hamiltonian structure. In Section 4, a second-order energy-conserving temporal scheme for quasi-Hamiltonian systems is presented. It resembles a Crank-Nicholson scheme or an implicit midpoint rule, and is implemented using a similar quasi-Newton iteration. In Section 5, we conduct numerical experiments to verify

[1] In fact, a quasi-Hamiltonian structure was a design principle for this discretization, but the presentation avoided this aspect in an effort to improve accessibility.

the properties of our scheme, such as: order of convergence and exact conservation of total mass, buoyancy and energy. Section 6 summarizes and discusses our results.

2. Continuous Equations

Let the domain $\Omega \subset \mathbb{M}$ be a compact, closed two dimensional subset of a manifold $\mathbb{M}$, that is rotating with some constant rotation rate Ω, which gives rise to a (possibly spatially varying) Coriolis parameter f. Boundary terms can be introduced in what follows in order to handle compact domains with boundaries, such a β channel, but this is deferred to later work. Useful operators are the gradient ∇, divergence $\nabla\cdot$, skew-gradient $\nabla^T = \hat{\mathbf{k}} \times \nabla$, curl $\nabla^T\cdot = \hat{\mathbf{k}} \cdot \nabla\times$ and vector transpose $\mathbf{x^T} = \hat{\mathbf{k}} \times \mathbf{x}$; where $\hat{\mathbf{k}}$ is the local unit vertical vector. This section is a review and compilation of existing material in the literature, and no claim to originality is made.

2.1. Thermal Shallow Water Equations

Now consider a thin layer of hydrostatic Boussinesq fluid in Ω, moving in columns, on top of a rigid bottom with height $b = b(x, y)$ and with horizontally varying density $\rho = \rho(x, y, t)$, where $\bar{\rho}$ is the density used in the Boussinesq approximation. The equations of motion that govern such a system are known as the thermal shallow water equations, and they are written in curl form [1] as:

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\,\mathbf{u}) = 0 \tag{1}$$

$$\frac{\partial\,\mathbf{u}}{\partial t} + hq\,\mathbf{u}^T + \nabla(\frac{\mathbf{u}\cdot\mathbf{u}}{2}) + s\nabla(h+b) + \frac{h}{2}\nabla s = 0 \tag{2}$$

$$\frac{\partial s}{\partial t} + \mathbf{u}\cdot\nabla s = 0 \tag{3}$$

Here h is the fluid height, $\mathbf{u}$ is the relative velocity, $s = g\frac{\rho}{\bar{\rho}}$ is the buoyancy, $S = hs$ is the mass-weighted buoyancy, g is the (constant) gravity, b is the topographic height, $\zeta = \nabla^T \cdot \mathbf{u}$ is the relative vorticity, $\eta = \zeta + f$ is the absolute vorticity and $q = \frac{\eta}{h}$ is the potential vorticity. The standard shallow water equations are recovered for the case of constant buoyancy $s = g$ i.e. $\rho = \bar{\rho}$. In the multilayer extension, a constant s in each layer would represent a barotropic flow, while allowing s to vary leads to baroclinic flows. Combining (1) and (3), an evolution equation for S can be written as

$$\frac{\partial S}{\partial t} + \nabla \cdot (sh\,\mathbf{u}) = 0 \tag{4}$$

It is also possible to rewrite the momentum equation (2) to involve ∇S instead of ∇s as

$$\frac{\partial\,\mathbf{u}}{\partial t} + hq\,\mathbf{u}^T + \nabla(\frac{S}{2}) + \nabla(\frac{\mathbf{u}\cdot\mathbf{u}}{2}) + s\nabla(\frac{h}{2} + b) = 0 \tag{5}$$

Two possible choice for prognostic variables are $(h, \mathbf{u}, S)$ and $(h, \mathbf{u}, s)$, which are discussed in the following sections.

Hamiltonian Formulation. As detailed in [1], the thermal shallow water equations in Ω can be described using a Hamiltonian $\mathcal{H}$ and a Poisson bracket, which is a bilinear, anti-symmetric operator that satisfies the Jacobi identity and the Leibniz rule. The dynamics of an arbitrary functional $\mathcal{F}[x]$ (where x are the prognostic variables) are then given by

$$\frac{d\,\mathcal{F}}{dt} = \{\mathcal{F}, \mathcal{H}\} \tag{6}$$

The time evolution of a particular prognostic variable x_i can be obtained by choosing $\mathcal{F} = \int \hat{x}_i x_i$, where $\hat{x}_i$ is an arbitrary function belonging to the same infinite-dimensional function space as x_i (also commonly referred to as a test function). This leads to the weak-form equations, from which the usual strong form equations are deduced upon the assumption of smoothness. Typically, for fluid dynamical systems in Eulerian coordinates the Poisson bracket is non-canonical: it has a non-trivial nullspace. There exists a set of functionals, termed Casimirs $\mathcal{C}$, that satisfy:

$$\{\mathcal{A}, \mathcal{C}\} = 0 \tag{7}$$

for any functional $\mathcal{A}$. Therefore $\frac{d\mathcal{C}}{dt} = \{\mathcal{C}, \mathcal{H}\} = 0$, and these quantities are conserved.

2.2. Predicting $(h, \mathbf{u}, S)$

The first approach predicts the mass-weighted buoyancy S.

Hamiltonian. The Hamiltonian $\mathcal{H}[h, \mathbf{u}, S]$ is given by

$$\mathcal{H}[h, \mathbf{u}, S] = \int_\Omega \frac{Sh}{2} + Sb + h\frac{\mathbf{u}\cdot\mathbf{u}}{2} \tag{8}$$

with associated functional derivatives

$$\frac{\delta\,\mathcal{H}}{\delta h} := B = \frac{S}{2} + \frac{\mathbf{u}\cdot\mathbf{u}}{2} \qquad \frac{\delta\,\mathcal{H}}{\delta\,\mathbf{u}} := \mathbf{F} = h\,\mathbf{u} \qquad \frac{\delta\,\mathcal{H}}{\delta S} := T = \frac{h}{2} + b \tag{9}$$

Poisson Bracket. The Poisson bracket is given by

$$\{\mathcal{A}, \mathcal{B}\} = \{\mathcal{A}, \mathcal{B}\}_R + \{\mathcal{A}, \mathcal{B}\}_Q + \{\mathcal{A}, \mathcal{B}\}_S \tag{10}$$

where:

$$\{\mathcal{A}, \mathcal{B}\}_R = \int_\Omega (-\frac{\delta\,\mathcal{A}}{\delta h}\nabla\cdot\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}} + \frac{\delta\,\mathcal{B}}{\delta h}\nabla\cdot\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}}) \tag{11}$$

$$\{\mathcal{A}, \mathcal{B}\}_Q = \int_\Omega -q\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}}\cdot\frac{\delta\,\mathcal{A}^T}{\delta\,\mathbf{u}} \tag{12}$$

$$\{\mathcal{A}, \mathcal{B}\}_S = \int_\Omega (-\frac{\delta\,\mathcal{A}}{\delta S}\nabla\cdot(s\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}}) + \frac{\delta\,\mathcal{B}}{\delta S}\nabla\cdot(s\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}})) \tag{13}$$

These brackets conserve $\mathcal{H}$ by virtue of their anti-symmetry, and since they are non-canonical there will be additional conserved quantities known as Casimirs (see below).

Equations of Motion. The equations of motion that come from the substituting the functional derivatives (9) into the Poisson brackets (11) - (13) are:

$$\frac{\partial h}{\partial t} + \nabla \cdot \mathbf{F} = 0 \tag{14}$$

$$\frac{\partial \mathbf{u}}{\partial t} + q\,\mathbf{F}^T + \nabla B + s\nabla T = 0 \tag{15}$$

$$\frac{\partial S}{\partial t} + \nabla \cdot (s\,\mathbf{F}) = 0 \tag{16}$$

Upon substitution of actual values of functional derivatives (9), it is easy to see that (14) - (16) are the same as (1) - (4). When $s = g$ the shallow water equations are recovered.

Linearized Equations. The dynamics associated with small-amplitude perturbations around a resting steady state $(h, \mathbf{u}, S) = (H, 0, gH)$ (with $b = 0$), where H is a constant, can be derived following well-established procedures in Hamiltonian mechanics [50]. This is commonly referred to as linearization, and proceeds as follows. The linearized Hamiltonian $\mathcal{H}_L[h, \mathbf{u}, S]$, which is the small-amplitude pseudo-energy associated with the steady state, is

$$\mathcal{H}_L[h, \mathbf{u}, S] = \int_\Omega \frac{Sh + ghH + SH}{2} + H\frac{\mathbf{u}\cdot\mathbf{u}}{2} \tag{17}$$

with associated functional derivatives

$$\frac{\delta\,\mathcal{H}_L}{\delta h} = \frac{S}{2} + \frac{gH}{2} \qquad \frac{\delta\,\mathcal{H}_L}{\delta\,\mathbf{u}} = H\,\mathbf{u} \qquad \frac{\delta\,\mathcal{H}_L}{\delta S} = \frac{h}{2} + \frac{H}{2} \tag{18}$$

The dynamics are then given by the Poisson brackets evaluated at the steady state. The $\{\mathcal{A}, \mathcal{B}\}_R$ bracket is unchanged, while $\{\mathcal{A}, \mathcal{B}\}_Q \to \{\mathcal{A}, \mathcal{B}\}_W$ and $\{\mathcal{A}, \mathcal{B}\}_S \to \{\mathcal{A}, \mathcal{B}\}_{SL}$ where

$$\{\mathcal{A}, \mathcal{B}\}_W = \int_\Omega -\frac{f}{H}\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}} \cdot \frac{\delta\,\mathcal{B}^T}{\delta\,\mathbf{u}} \tag{19}$$

$$\{\mathcal{A}, \mathcal{B}\}_{SL} = \int_\Omega (-\frac{\delta\,\mathcal{A}}{\delta S}\nabla \cdot (g\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}}) - g\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}} \cdot \nabla\frac{\delta\,\mathcal{B}}{\delta S}) \tag{20}$$

Substituting functional derivatives (18) into the Poisson brackets (11) and (19) - (20) gives the equations of motion as

$$\frac{\partial h}{\partial t} + H\nabla \cdot \mathbf{u} = 0 \tag{21}$$

$$\frac{\partial \mathbf{u}}{\partial t} + f\,\mathbf{u}^T + \nabla\frac{S}{2} + g\nabla\frac{h}{2} = 0 \tag{22}$$

$$\frac{\partial S}{\partial t} + gH\nabla \cdot \mathbf{u} = 0 \tag{23}$$

When $S = gh$, we recover the linearized shallow water equations. This is not a particularly physically interesting linearization (there are much better found in [51]), but it suffices as the basis for a simplified Jacobian to be used in solving the nonlinear system that arises from an implicit time stepping scheme, which is what it is needed for.

Casimirs. The Casimirs are given by

$$\mathcal{C}[h, \mathbf{u}, S] = \int_\Omega hqG(\frac{S}{h}) + hK(\frac{S}{h}) \tag{24}$$

where G and K are arbitrary functions of s. The associated functional derivatives are

$$\frac{\delta \mathcal{C}}{\delta h} = K - sqG' - sK' \qquad \frac{\delta \mathcal{C}}{\delta \mathbf{u}} = -\nabla^T G \qquad \frac{\delta \mathcal{C}}{\delta S} = qG' + K' \tag{25}$$

Important cases are total mass ($G = 0$,$K = 1$), total potential vorticity ($G = 1$,$K = 0$) and total buoyancy ($G = 0$,$K = s$). Unlike the shallow water equations, potential enstrophy and higher moments of the potential vorticity are no longer conserved. This is because potential vorticity is no longer a material invariant. Taking $\nabla\cdot$ of (15) gives the vorticity equation

$$\frac{\partial \eta}{\partial t} + \nabla \cdot (hq\,\mathbf{u}) + \nabla^T \cdot \left(s\nabla(\frac{h}{2} + b) \right) = 0 \tag{26}$$

which yields (noting $\eta = hq$)

$$\frac{Dq}{Dt} + \frac{1}{h}\nabla^T \cdot \left(s\nabla(\frac{h}{2} + b) \right) = 0 \tag{27}$$

The extra terms (compared to the shallow water equations, which have $\frac{Dq}{Dt} = 0$) are zero only when s is a constant.

2.3. Predicting $(h, \mathbf{u}, s)$

It is also possible to predict buoyancy s instead of mass-weighted buoyancy S. Using the chain rule, the functional derivatives of an arbitrary functional $\mathcal{A}'[h, \mathbf{u}, s]$ in terms of those for $\mathcal{A}[h, \mathbf{u}, S]$ can be written as

$$\frac{\delta \mathcal{A}'}{\delta h} = \frac{\delta \mathcal{A}}{\delta h} + s\frac{\delta \mathcal{A}}{\delta S} \qquad \frac{\delta \mathcal{A}'}{\delta \mathbf{u}} = \frac{\delta \mathcal{A}}{\delta \mathbf{u}} \qquad \frac{\delta \mathcal{A}'}{\delta S} = h\frac{\delta \mathcal{A}}{\delta S} \tag{28}$$

Hamiltonian. The Hamiltonian is

$$\mathcal{H}'[h, \mathbf{u}, s] = \int_\Omega \frac{h^2 s}{2} + hsb + h\frac{\mathbf{u}\cdot\mathbf{u}}{2} \tag{29}$$

which gives

$$\frac{\delta \mathcal{H}'}{\delta h} := B' = \frac{\mathbf{u}\cdot\mathbf{u}}{2} + sh + sb \qquad \frac{\delta \mathcal{H}'}{\delta \mathbf{u}} := \mathbf{F} = h\,\mathbf{u} \qquad \frac{\delta \mathcal{H}'}{\delta s} := T' = \frac{1}{2}h^2 + hb \tag{30}$$

These can be obtained either by using the chain rule (28) in (9), or by directly taking functional derivatives of (29).

Poisson Bracket. The chain rule (28) is also used to transform the brackets (11) - (13). They become

$$\{\mathcal{A}', \mathcal{B}'\}_R = \int_\Omega (-\frac{\delta \mathcal{A}'}{\delta h} \nabla \cdot \frac{\delta \mathcal{B}'}{\delta \mathbf{u}} + \frac{\delta \mathcal{B}'}{\delta h} \nabla \cdot \frac{\delta \mathcal{B}'}{\delta \mathbf{u}}) \tag{31}$$

$$\{\mathcal{A}', \mathcal{B}'\}_Q = \int_\Omega -q \frac{\delta \mathcal{A}'}{\delta \mathbf{u}} \cdot \frac{\delta \mathcal{B}'^T}{\delta \mathbf{u}} \tag{32}$$

$$\{\mathcal{A}', \mathcal{B}'\}_S = \int_\Omega \frac{\nabla s}{h} \cdot \left(\frac{\delta \mathcal{A}'}{\delta \mathbf{u}} \frac{\delta \mathcal{B}'}{\delta s} - \frac{\delta \mathcal{B}'}{\delta \mathbf{u}} \frac{\delta \mathcal{A}'}{\delta s} \right) d\Omega \tag{33}$$

Note that (31) - (32) have the same form as (11) - (12).

Equations of Motion. Putting the functional derivatives (30) into the Poisson brackets (31) - (33) gives the equations of motion as

$$\frac{\partial h}{\partial t} + \nabla \cdot \mathbf{F} = 0 \tag{34}$$

$$\frac{\partial \mathbf{u}}{\partial t} + q \mathbf{F}^T + \nabla B' - \frac{T'}{h} \nabla s = 0 \tag{35}$$

$$\frac{\partial s}{\partial t} + \frac{\mathbf{F}}{h} \cdot \nabla s = 0 \tag{36}$$

Again, substitution of the actual values for (30) into (34) - (36) yields (1) - (4), and when $s = g$ is the shallow water equations are recovered.

Linearized Equations. Linearizing about the same state (which has $s = g$), the chain rule between functional derivatives becomes

$$\frac{\delta \mathcal{A}'}{\delta h} = \frac{\delta \mathcal{A}}{\delta h} + g \frac{\delta \mathcal{A}}{\delta S} \qquad \frac{\delta \mathcal{A}'}{\delta \mathbf{u}} = \frac{\delta \mathcal{A}}{\delta \mathbf{u}} \qquad \frac{\delta \mathcal{A}'}{\delta S} = H \frac{\delta \mathcal{A}}{\delta S} \tag{37}$$

This is understood by noting that the linearized version of the relationship $S = sh$ is

$$S = gh + sH \tag{38}$$

Therefore the functional derivatives of the linearized Hamiltonian $\mathcal{H}'_L[h, \mathbf{u}, s]$

$$\mathcal{H}'_L[h, \mathbf{u}, s] = \int_\Omega \frac{shH}{2} + g\frac{h^2}{2} + ghH + \frac{sH^2}{2} + H\frac{\mathbf{u} \cdot \mathbf{u}}{2} \tag{39}$$

are

$$\frac{\delta \mathcal{H}'_L}{\delta h} = gh + \frac{sH}{2} + gH \qquad \frac{\delta \mathcal{H}'_L}{\delta \mathbf{u}} = H\,\mathbf{u} \qquad \frac{\delta \mathcal{H}'_L}{\delta s} = \frac{hH}{2} + \frac{H^2}{2} \tag{40}$$

The $\{\mathcal{A}', \mathcal{B}'\}_R$ bracket is unchanged, while $\{\mathcal{A}', \mathcal{B}'\}_Q \to \{\mathcal{A}', \mathcal{B}'\}_W = \{\mathcal{A}, \mathcal{B}\}_W$ and $\{\mathcal{A}', \mathcal{B}'\}_s \to \{\mathcal{A}', \mathcal{B}'\}_{SL} = 0$. This exposes the fact that $\frac{\partial s}{\partial t} = 0$ for this choice of linearization. This can be seen in the S variant, but it is somewhat hidden. The equations of

motion are

$$\frac{\partial h}{\partial t} + H\nabla \cdot \mathbf{u} = 0 \tag{41}$$

$$\frac{\partial \mathbf{u}}{\partial t} + f\,\mathbf{u}^T + g\nabla h + \frac{H}{2}\nabla s = 0 \tag{42}$$

$$\frac{\partial s}{\partial t} = 0 \tag{43}$$

As before, when $s = 0$ we recover the linear shallow water equations.

Casimirs. If s is predicted, the Casimirs $\mathcal{C}'[h, \mathbf{u}, s] = \mathcal{C}[h, \mathbf{u}, S]$ are written in the form

$$\mathcal{C}'[h, \mathbf{u}, s] = \int_\Omega hqG(s) + hK(s) \tag{44}$$

and by the chain rule (28) or direct calculation we have

$$\frac{\delta\,\mathcal{C}'}{\delta h} = K \qquad \frac{\delta\,\mathcal{C}'}{\delta\,\mathbf{u}} = -\nabla^T G \qquad \frac{\delta\,\mathcal{C}'}{\delta s} = hqG' + hK' \tag{45}$$

3. Compatible Galerkin Spatial Discretization

The thermal shallow water equations (14) - (16) or (34) - (36) are discretized following a Galerkin approach. Utilizing the Hamiltonian formulation, a particularly elegant procedure is

1. Restrict the Poisson bracket (and possibly perform integration by parts) to a set of compatible Galerkin spaces, which is one that forms a discrete version of the deRham complex [52, 53]. The construction of such spaces is described in Section 3.2.

2. Discretize the functionals $\mathcal{H}$ and $\mathcal{F}$ using variables from the same spaces.

Then the discrete equations of motion come from simply inserting the discrete functional derivatives into the discretized brackets. This is the Galerkin version of the finite-difference approach advocated in [43] (and further developed in [2, 42, 45, 46, 47]) which discretizes the functionals $\mathcal{H}$ and $\mathcal{F}$, and the Poisson bracket $\{\mathcal{A}, \mathcal{B}\}$ directly, rather than the equations of motion themselves. Using compatible Galerkin spaces ensures that the discrete brackets are anti-symmetric and have some subset of the Casimirs of the continuous bracket. However, it does not seem possible to also give the discrete brackets the Jacobi identity. We therefore refer to this approach as a quasi-Hamiltonian discretization. Before pursing this idea, first a set of desirable properties for a numerical model of the thermal shallow water equations to possess is identified.

3.1. Desirable Properties

As discussed in [54], there are many desirable properties that a numerical model of the atmosphere should posses. Here we propose a similar list for the case of the thermal shallow water equations, and give a general approach to achieve them. We separate the properties into two groups:

Category 1

(A1) Conservation of total mass through a flux based formulation

(A2) Conservation of total buoyancy, through a flux based formulation

(A3) Conservation of total energy, through the correct energy conversion terms (note that this also implies the linearized equations conserve energy):

- (a) The pressure gradient terms ∇B and $\nabla \cdot \mathbf{F}$ cancel to conserve energy (equivalent to an anti-symmetric $\{\mathcal{A}, \mathcal{B}\}_R = \{\mathcal{A}', \mathcal{B}'\}_R$ bracket)
- (b) The Coriolis term $q\,\mathbf{F}^T$ is energy conserving (equivalent to an anti-symmetric $\{\mathcal{A}, \mathcal{B}\}_Q = \{\mathcal{A}', \mathcal{B}'\}_Q$ bracket)
- (c) The buoyancy gradient terms $s\nabla T$ and $\nabla \cdot (s\,\mathbf{F})$ cancel to conserve energy (or their counterparts in the s variant; equivalent to an anti-symmetric $\{\mathcal{A}, \mathcal{B}\}_S$ or $\{\mathcal{A}', \mathcal{B}'\}_s$ bracket)

(A4) Conservation of total potential vorticity $\int hq = \int \eta$, through

- (a) The pressure gradient term ∇B does not produce vorticity
- (b) The entropy gradient term $s\nabla T$ (or the counterpart in the s variant) does not produce vorticity

(A5) Compatible advection of s: If s is uniform initially, then $\frac{\partial s}{\partial t} = 0$. If predicting S, this is equivalent to the S equation becoming the h equation when $s = 1$.

(A6) At least 2nd order Taylor series accuracy (preferably arbitrary order)

Additionally, the implied shallow water discretization (that is obtained when $s = g$) should satisfy:

(B1) Conservation of mass through a flux based formulation

(B2) Conservation of total energy, through the correct energy conversion terms (note that this also implies the linearized equations conserve energy):

- (a) The pressure gradient terms ∇B and $\nabla \cdot \mathbf{F}$ cancel to conserve energy (equivalent to an anti-symmetric $\{\mathcal{A}, \mathcal{B}\}_R = \{\mathcal{A}', \mathcal{B}'\}_R$ bracket)

(b) The Coriolis term $q\,\mathbf{F}^T$ is energy conserving (equivalent to an anti-symmetric $\{\mathcal{A},\mathcal{B}\}_Q = \{\mathcal{A}',\mathcal{B}'\}_Q$ bracket)

(B3) Conservation of total potential vorticity, through a flux based formulation

(B4) Conservation of potential enstrophy, through the correct conversion terms

(B5) Compatible potential vorticity advection for the nonlinear equations = Steady geostrophic modes for the linear equations: A initially uniform q field should remain uniform

(B6) Good linear mode properties:

(a) Free of spurious stationary modes, such as pressure modes

(b) Free of spurious inertial modes

Note that generally, (A1), (A3) and (A4) will imply (B1), (B2) and (B3), respectively. Spurious stationary modes are non-propagating linear modes that do not have a physical analogue, and are often damaging to simulations. A prominent example is the pressure mode that arises for unstabilized P_n - P_n discretizations of the shallow water equations. Spurious inertial modes are propagating but not wavelike, unphysical linear modes related to the physical inertial mode. They occur in P_m - P_n ($m > n$) discretizations of the shallow water equations. More discussion of spurious stationary and inertial modes can be found in [30]. Non-steady geostrophic modes are known to arise for certain hexagonal C-grid finite difference models (also hexagonal B/D/E grids) [33] (where they spuriously propagate but do not decay) and for discontinuous Galerkin methods (Daniel Le Roux, personal communication, where they decay but do not propagate).

Category 2

(C1) Good linear mode properties for the implied shallow water discretization:

(a) Free of spurious branches of dispersion relationship

(b) Discrete dispersion relationships that are good approximations of the continuous ones, without spectral gaps and with higher-order accuracy

(C2) Computational efficiency on a range of architectures, including expected future trends towards high levels of parallelism and increasing memory hierarchy depths

(C3) Freedom from the Hollingsworth instability

Spurious branches of the dispersion relationship are unphysical inertia-gravity or Rossby wave branches, that arise due to a mismatch in the number of degrees of freedom (finite-difference models) or dimensionality of the spaces (finite element models) between the wind and mass fields. This typically occurs on non-quadrilateral grids, although they also occur for the $S_r\Lambda^k$ family on quadrilaterals from finite element exterior calculus. Spectral gaps are

non-dimensional wavenumbers where the dispersion relationship is double-valued and the group velocity goes to zero, and they are unphysical artifacts that lead to noisy simulations and incorrect propagation of wave packets with energy at the gap frequencies. They often occur for higher-order finite element methods. The Hollingsworth instability [38, 39] is a nonlinear numerical instability seen in 3D models for small equivalent depths (slow internal modes), typically for C grid finite difference models using energy-conserving discretizations. It is though to arise from a mismatch between the advective and vector invariant forms of momentum transport, and despite several decades passing since its discovery it is still not well understood. However, recently some progress has been made [37, 40, 41].

How to obtain them. This is a rather long list of properties, but they are all achievable through using compatible Galerkin methods to construct a quasi-Hamiltonian discretization. A compatible Galerkin space discretization is a generalization of finite element exterior calculus [52, 55] to any set of spaces that construct a discrete deRham complex. There are many known sets of spaces that do this, for examples see [9, 53]. The specific quasi-Hamiltonian discretizations presented in Section 3.5, 3.6 and 3.7 achieve (A1) - (A5), which is demonstrated in Section 3.8, 3.9 and 3.10; and they are arbitrary-order, giving (A6). As demonstrated in [9, 10, 11], a quasi-Hamiltonian discretization of the shallow water equations using compatible Galerkin methods gives (B1) - (B6), which the discretizations in this paper reduce to when $s = g$.

The specific choice of compatible Galerkin spaces and their implementation then determines the ability to satisfy the properties (C1) and (C2) in Category 2, and Table 1 details the ability of various choices of compatible Galerkin families to obtain properties (C1) and (C2). In particular, on triangular meshes the standard finite element exterior calculus families $P_r^-\Lambda^k$ and $P_r\Lambda^k$ have spurious branches [9], spectral gaps and there is not a known tensor-product formulation for the basis functions. It is possible to avoid spurious branches on triangles using the $BDFM_1$ space [9], but this still does not avoid spectral gaps and the extension of this approach to higher-order is unclear. For these reasons, we have chosen instead to use tensor-product Galerkin methods on structured, quadrilateral meshes. In this approach, a set of compatible Galerkin spaces in 1D are chosen, and tensor products are used to construct the 2D spaces. More details are found in Section 3.2. A tensor-product structure enables many computational techniques such as sum factorization that accelerate the resulting code, helping satisfy (C2). Additionally, a tensor-product structure ensures that spurious branches of the dispersion relationship are avoided at any order of accuracy, since there are always two degrees of freedom in the wind field for every degree of freedom in the mass field; this gives (C1a). However, the 1D spaces must be chosen with care. As shown in [26, 27, 31], the standard finite element exterior calculus choice of $Q_r^-\Lambda^k$ in 2D and $P_n^C - P_{n-1}^{DG}$ in 1D, where P_n^C is the order n Lagrange space and P_{n-1}^{DG} is the order $n-1$ discontinuous Lagrange space, leads to the presence of spectral gaps when $n \geq 2$. A set of spaces, called mimetic Galerkin differences (MGD_n), that avoids spectral gaps are introduced in [24, 49] and discussed further below.

It should be noted that all of the compatible Galerkin methods investigated in the literature and by the authors suffer from a CD mode [30] (also known as the f-mode or Coriolis

Element family	C1a Spurious Branches	C1b Spectral Gaps	C2 Tensor-Product Structure
$P_r^-\Lambda^k$	Yes[1]	Likely[3]	No
$P_r\Lambda^k$	Yes[2]	Likely[3]	No
$S_r\Lambda^k$	Yes[2]	Likely[3]	No
$BDFM_1$	No	Likely[3]	No
$Q_r^-\Lambda^k$	No	Yes[4]	Yes
MGD_n	No	No	Yes

[1]: Spurious inertia-gravity waves [2]: Spurious Rossby waves
[3]: Proven for $1D$ version ($P_n^C - P_{n-1}^{DG}$), seems extremely likely for $2D$
[4]: Lowest-order avoids gaps

Table 1: Effects of space choice on ability to obtain the desirable properties of Section 3.1. All spaces achieve the Category 1 properties. Here we have used the finite element exterior calculus [55, 52] names for the families, where appropriate. Only the MGD_n spaces are able to obtain all of the properties. The $P_r^-\Lambda^k$ (RT_n on triangles for velocity), $P_r\Lambda^k$ (BDM_n on triangles for velocity), $Q_r^-\Lambda^k$ (RT_n on quadrilaterals for velocity) and $BDFM_1$ families were investigated in [9], where the presence or lack of spurious branches was demonstrated. The presence of spurious Rossby modes might be acceptable (as in hexagonal C grid finite difference schemes), but spurious inertia-gravity modes (such as those present for the triangular C grid) are known to cause issues [56]. The $S_r\Lambda^k$ family has never been studied (to the author's knowledge) in geophysical fluid dynamics before, but the analysis in [9] applies and indicates that it will have spurious Rossby modes (too many wind degrees of freedom compared to mass).

mode), which occurs due to the rank deficiency of the discrete Coriolis matrix. The CD mode is the zero group velocity mode occurring at the highest wavenumber associated with the averaging required for the Coriolis term. It appears to be an unavoidable feature of compatible Galerkin (and C-grid finite difference) methods. This mode does not appear to have any detrimental impact on simulations with a sufficiently resolved Rossby radius, which is the case for most realistic atmospheric models today.

A detailed study of the Hollingsworth instability for compatible Galerkin methods is still lacking. However, we have run the test case discussed in Section 5.2 of [37] with $R_u = 20$ and $F_u = 10$ using the EC2-SI time integrator and MGD_3 family, and found no sign of the Hollingsworth instability out to a non-dimensional time of $T = 4$. Similar results were obtained for a different but related discretization using the BDM_2 family on triangles (Pedro Peixoto, personal communication). We are currently undertaking a numerical version of the stability study performed in [37] for several compatible Galerkin families on triangles and quadrilaterals (including MGD_n) and discretization choices for the nonlinear PV flux term, and will report on the results of this in a future publication.

In summary, we have chosen to discretize the Hamiltonian formulation using a tensor-product compatible Galerkin method, with a specific choice of underlying 1D spaces such that a good dispersion relationship is obtained for arbitrary order. In this way, our spatial discretization is able to obtain all of the desired properties. To our knowledge, this is the first discretization that achieves this.

3.2. Mimetic Galerkin Differences (MGD_n)

Start with a set of 1D spaces $\mathbb{A} \subset H^1$ and $\mathbb{B} \subset L_2$ that are a partition of unity and form the 1D discrete deRham complex $\mathbb{A} \xrightarrow{\partial/\partial x} \mathbb{B}$. From these it is possible to construct a 2D discrete deRham complex satisfying the necessary properties by forming the subspaces of H^1, L_2 and $H(div)$ using tensor products of the 1D spaces:

$$\mathbb{W}_0 = \mathbb{A} \otimes \mathbb{A} \subset H^1 \tag{46}$$

$$\mathbb{W}_1 = (\mathbb{A} \otimes \mathbb{B})\hat{\mathbf{i}} + (\mathbb{B} \otimes \mathbb{A})\hat{\mathbf{j}} \subset H(div) \tag{47}$$

$$\mathbb{W}_2 = \mathbb{B} \otimes \mathbb{B} \subset L_2 \tag{48}$$

where $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ are unit vectors on the reference quadrilateral in the $\hat{x}$ and $\hat{y}$ directions, respectively. More details about the tensor product construction can be found in [11, 53, 57]. One possible choice is $\mathbb{A} = P_n^C$ and $\mathbb{B} = P_{n-1}^{DG}$, which gives rise to the $Q_r^-\Lambda^k$ family from finite element exterior calculus. Alternative choices give rise to the mimetic spectral element method [53] and isogeometric discrete differential forms [53, 58, 59]. We choose instead to use Galerkin differences [60] GD_n for $\mathbb{A}$, and to define the corresponding space $\mathbb{B} = DGD_{n-1}$ (referred to as discontinuous Galerkin differences) using the ideas in [53]. We will refer to this approach as the mimetic Galerkin differences (MGD_n) family, which is defined for n odd. Specifically, consider the m basis functions $N_i(x)$ for GD_n, where $i \in [0, \ldots, m-1]$, on a periodic 1D mesh with m elements. The corresponding m basis functions $M_j(x)$ (with $j \in [1, \ldots, m]$) for DGD_{n-1} are given by

$$M_j(x) = -\sum_{k=0}^{j-1} \frac{d}{dx} N_k(x) = \sum_{k=j}^{m-1} \frac{d}{dx} N_k(x) \tag{49}$$

where the two expressions are identical since $\mathbb{A} = GD_n$ forms a partition of unity and therefore

$$\sum_{i=0}^{m-1} N_i(x) = 1 \rightarrow \sum_{i=0}^{m-1} \frac{d}{dx} N_i(x) = 0 \tag{50}$$

There are the same number of degrees of freedom for $\mathbb{A}$ and $\mathbb{B}$ since the domain is periodic. Figure 1 shows plots of the basis functions for the $\mathbb{W}_0$, $\mathbb{W}_1$ and $\mathbb{W}_2$ spaces of the MGD_3 family. Note that unlike finite elements, the basis function for a degree of freedom has support that is not limited to the topological support of the associated geometric entity, although the support is still compact. As a historical note, the $n = 3$ version of these elements were developed using different methods, discussed further in [49]. Following standard practice for element based Galerkin methods, integrals are evaluated first by pulling back to a reference element (using the appropriate Piola transform), and then calculating the resulting transformed integral numerically with a quadrature rule of the appropriate degree. More details on this are found in Section 3.4. Optimal quadrature rules for the standard $Q_r^-\Lambda^k$ family are known (they are simply Gaussian quadrature for each element), but optimal quadrature rules for

other compatible Galerkin families, including the MGD_n family, remain a subject of active research (see [61] for an example of such rules for isogeometric analysis).

As detailed in [24], the $GD_n - DGD_{n-1}$ pair in $1D$ gives an inertia-gravity wave dispersion relationship that is free of spectral gaps for any n. This is in contrast to $P_n^C - P_{n-1}^{DG}$ element, which has spectral gaps [26] for $n \geq 2$. Although it is possible to eliminate these gaps for $n = 2$ through partial lumping of the velocity mass matrix, it does not seem possible to do so for $n \geq 3$. In any case, mass lumping is an equation dependent procedure that also destroys the higher-order convergence of the dispersion relationship. This is a strong motivation for the use of mimetic Galerkin differences. Numerical calculations have confirmed that the spectral gaps are also eliminated in the case of the 2D linear shallow water equations, and work is ongoing to verify this analytically.

One major advantage of these spaces is that there is a single degree of freedom per geometric entity (vertex, edge or cell) of the mesh. In fact, the basis coefficients at any order n are scalars sampled at vertices for the $\mathbb{W}_0$ space, fluxes integrated over edges for the $\mathbb{W}_1$ space; and densities integrated over cells for the $\mathbb{W}_2$ space. These are precisely the degrees of freedom for a standard C grid finite-difference method, and this is expected to make coupling to existing physics parameterizations and tracer transport schemes for the development of a full model much simpler than with other Galerkin methods. This property is also shared with the lowest-order $Q_r^- \Lambda^k$ family (corresponding to the $P_1^C - P_0^{DG}$ element in 1D) for an appropriate choice of basis function, but not any of the higher order variants. The mimetic spectral element method also shares this property, albeit for a non-uniform subelement set of cells, vertices and edges [19]. The MGD_n family is shown schematically Figure 2.

3.3. Prognostic Variables and Choice of Spaces

The discretizations presented in this section predict h and $\mathbf{u}$, and either S or s. If S is predicted, s must be diagnosed. Additionally, if the indirect form of the nonlinear PV flux term is used, then q must also be diagnosed. Following ideas from differential geometry [62, 63, 64], we place $h, S, b, B, B' \in \mathbb{W}_2$, $\mathbf{u}, \mathbf{F} \in \mathbb{W}_1$ and $q \in \mathbb{W}_0$. This corresponds with h and S being 2-forms, $\mathbf{u}$ and $\mathbf{F}$ being 1-forms and q being a 0-form. The test functions (denoted with hats) are $\hat{h}, \hat{S} \in \mathbb{W}_2$; $\hat{\mathbf{u}} \in \mathbb{W}_1$ and $\hat{q} \in \mathbb{W}_0$. For the variant that predicts S (Section 3.5), we also place $s, \hat{s} \in \mathbb{W}_2$. For the variant that predicts s (Section 3.6), we either place $s, \hat{s}, T' \in \mathbb{W}_2$ or $s, \hat{s}, T', h' \in \mathbb{W}_0$. Finally, it is possible to directly discretize the nonlinear PV flux term, instead of using the form involving q (Section 3.7). This amounts to discretizing a $\{\mathcal{A}, \mathcal{B}\}_Q$ bracket of the form

$$\{\mathcal{A}, \mathcal{B}\}_Q = \int_\Omega -\frac{\nabla \cdot \mathbf{u}}{h} \frac{\delta \mathcal{A}}{\delta \mathbf{u}} \cdot \frac{\delta \mathcal{B}^T}{\delta \mathbf{u}} \tag{51}$$

instead of (12). Then purely for diagnostic purposes it is useful to introduce absolute vorticity $\eta \in \mathbb{W}_0$ with $\hat{\eta} \in \mathbb{W}_0$. All of these possible choices are shown schematically in Table 2. This choice of spaces corresponds to an Arakawa C-Grid for a finite difference model. Table 3 shows the form of conserved quantities and other properties achieved by the various discretization variants presented in this section.

Variant	$\mathbb{W}_0$	$\mathbb{W}_1$	$\mathbb{W}_2$
Predict S, use q form	q	$\mathbf{u}$,$\mathbf{F}$	h,S,b,B,T,s
Predict S, use direct form	η	$\mathbf{u}$,$\mathbf{F}$	h,S,b,B,T,s
Predict $s \in \mathbb{W}_0$, use q form	q,h',s,T'	$\mathbf{u}$,$\mathbf{F}$	h,b,B'
Predict $s \in \mathbb{W}_0$, use direct form	η,h',s,T'	$\mathbf{u}$,$\mathbf{F}$	h,b,B'
Predict $s \in \mathbb{W}_2$, use q form	q	$\mathbf{u}$,$\mathbf{F}$	h,s,b,B',T'
Predict $s \in \mathbb{W}_2$, use direct form	η	$\mathbf{u}$,$\mathbf{F}$	h,s,b,B',T'

Table 2: Choice of spaces for prognostic variables, constants, diagnostic variables and auxiliary variables; for the 6 possible variants presented in this work. Prognostic variables (red) have a time evolution equation. Constants (light blue) do not change in time, and are set once at the beginning of the simulation. Diagnostic variables are determined from the prognostic variables as needed; and are divided into two categories: those associated with the Poisson bracket (blue) and those associated with the functional derivatives of the Hamiltonian (green). This distinction is useful when considering the time discretization, since the integrator treats the two types separately. Auxiliary variables (black) are useful for computing statistics and making plots, they are not needed for the evolution of the system.

Variant	A4 $\mathcal{H}$	A2 $\mathcal{B}$	A3 and B1 $\mathcal{PV}$	B2 $\mathcal{PE}$	B3 PV compatability
Predict S, use q form	$\frac{1}{2}\langle S, h+2b\rangle + \mathcal{H}_K$	$\langle 1, S\rangle$	$\langle h, q\rangle$	Yes	Yes
Predict S, use direct form	$\frac{1}{2}\langle S, h+2b\rangle + \mathcal{H}_K$	$\langle 1, S\rangle$	$\langle 1, \eta\rangle$	No	No
Predict $s \in \mathbb{W}_0$, use q form	$\frac{1}{2}\langle hs, h+2b\rangle + \mathcal{H}_K$	$\langle h', s\rangle$	$\langle h, q\rangle$	Yes	Yes
Predict $s \in \mathbb{W}_0$, use direct form	$\frac{1}{2}\langle hs, h+2b\rangle + \mathcal{H}_K$	$\langle h', s\rangle$	$\langle 1, \eta\rangle$	No	No
Predict $s \in L^2$, use q form	$\frac{1}{2}\langle hs, h+2b\rangle + \mathcal{H}_K$	$\langle h, s\rangle$	$\langle h, q\rangle$	Yes	Yes
Predict $s \in L^2$, use direct form	$\frac{1}{2}\langle hs, h+2b\rangle + \mathcal{H}_K$	$\langle h, s\rangle$	$\langle 1, \eta\rangle$	No	No

$\mathcal{H}_K = \frac{1}{2}\langle h\,\mathbf{u}, \mathbf{u}\rangle = \frac{1}{2}\langle \mathbf{F}, \mathbf{u}\rangle = \langle h, K\rangle$ with $K = \frac{1}{2}\,\mathbf{u}\cdot\mathbf{u}$

Table 3: Form of the conserved quantities total energy $\mathcal{H}$, total buoyancy $\mathcal{B}$ and total potential vorticity $\mathcal{PV}$ for the six variants discussed in the text; and their ability to obtain potential enstrophy $\mathcal{PE} = \frac{1}{2}\langle hq, q\rangle$ conservation and PV compatibility for the implied shallow water discretization in the limit $s = g$. Note that although all discretizations conserve discrete analogues of $\mathcal{H}$, $\mathcal{B}$ and $\mathcal{PV}$, these analogues are different. Unless otherwise noted, the variants achieve all of the other properties from Section 3.1.

3.4. Discrete Grid and Operators

The domain Ω is discretized using a conforming mesh of curvilinear quadrilateral elements e and facets (curvilinear lines) f, with facet normals denoted by $\hat{\mathbf{n}}$. Given a variable x, x^+ and x^- are the restrictions of x to each side of the facet, arbitrarily labeled $+$ and $-$. All calculations are done on a reference element, with Piola transformations used to pullback from physical space to reference space. The transformation $\mathbf{x} = F(\hat{\mathbf{x}})$ from reference space to Ω, where $\mathbf{x} = (x, y)$ are coordinates in Ω and $\hat{\mathbf{x}} = (\hat{x}, \hat{y})$ are coordinates in reference space, has associated Jacobian $\mathbf{J}$

$$\mathbf{J} = \begin{pmatrix} \frac{\partial x}{\partial \hat{x}} & \frac{\partial x}{\partial \hat{y}} \\ \frac{\partial y}{\partial \hat{x}} & \frac{\partial y}{\partial \hat{y}} \end{pmatrix} \tag{52}$$

The Piola transforms are then defined as

$$\beta = \hat{\beta} \qquad \text{for } \beta, \hat{\beta} \in \mathbb{W}_0 \tag{53}$$

$$\boldsymbol{\gamma} = \mathbf{J}^{-T}\hat{\boldsymbol{\gamma}} \qquad \text{for } \boldsymbol{\gamma}, \hat{\boldsymbol{\gamma}} \in \mathbb{W}_1 \tag{54}$$

$$\beta = \frac{\hat{\beta}}{|\mathbf{J}|} \qquad \text{for } \beta, \hat{\beta} \in \mathbb{W}_2 \tag{55}$$

where β is a scalar quantity in physical space, $\hat{\beta}$ a scalar quantity in reference space, $\boldsymbol{\gamma}$ a vector quantity in physical space and $\hat{\boldsymbol{\gamma}}$ a vector quantity in reference space. The situation is somewhat more complicated on a manifold, but provided the manifold is orientable all the needed transformations exist (see [65] for more information).

The L_2 inner products over elements $\langle a, b\rangle$ and over facets $\langle a, b\rangle_\Gamma$ are defined as

$$\langle a, b\rangle = \sum_e \int_e a \quad \langle a, b\rangle_\Gamma = \sum_f \int_f ab \tag{56}$$

where ab is the appropriate scalar product (for example, it is the dot product for vectors). Due to the choice of spaces (i.e. the discrete deRham complex), the divergence $\nabla^T\cdot$ for $\mathbb{W}_1$ and skew-gradient $\nabla\cdot$ for $\mathbb{W}_0$ are strong operators. However, the gradient ∇ for $\mathbb{W}_2$ and curl $\nabla^T\cdot$ for $\mathbb{W}_1$ are defined weakly as

$$\left\langle \hat{\mathbf{u}}, \tilde{\nabla}\phi \right\rangle = -\left\langle \nabla\cdot\hat{\mathbf{u}}, \phi \right\rangle \tag{57}$$

$$\left\langle \hat{q}, \tilde{\nabla}^T\cdot\mathbf{u} \right\rangle = -\left\langle \nabla^T\hat{q}, \mathbf{u} \right\rangle \tag{58}$$

where $\hat{\mathbf{u}} \in \mathbb{W}_1$ and $\hat{q} \in \mathbb{W}_0$ are arbitrary test functions. The gradient ∇ for $\mathbb{W}_0$ is strong, but it actually maps into a subspace $\widehat{\mathbb{W}_1} \subset H(curl)$ belonging to the other 2D discrete deRham complex that has strong ∇ for $\mathbb{W}_0$ and $\nabla^T\cdot$ for $\widehat{\mathbb{W}_1}$ operators and weak $\tilde{\nabla}^T$ for $\mathbb{W}_2$ and $\tilde{\nabla}\cdot$ for $\widehat{\mathbb{W}_1}$ operators. Given a vector $\mathbf{x} = (\mathbf{a}, \mathbf{b})$, the transpose is defined as

$$\mathbf{x}^\mathbf{T} = (-\mathbf{b}, \mathbf{a}) \tag{59}$$

For facet integrals it is useful to introduce the broken operators ∇_H, ∇_H^T, $\nabla_H \cdot$ and $\nabla_H^T \cdot$ that are local to an element. The jump operator $[x]$ is defined as

$$[x] = x^+ - x^- \tag{60}$$

for scalar and vector x. It is also useful to define a jump operator for vectors that produces a scalar as

$$[\mathbf{x}, \hat{\mathbf{n}}] = \mathbf{x}^+ \, \hat{\mathbf{n}}^+ + \mathbf{x}^- \, \hat{\mathbf{n}}^- \tag{61}$$

and a jump operator for scalars that produces a vector as

$$[x, \hat{\mathbf{n}}] = x^+ \, \hat{\mathbf{n}}^+ + x^- \, \hat{\mathbf{n}}^- \tag{62}$$

The average operator $\{x\}$ is defined as

$$\{x\} = \frac{1}{2}(x^+ + x^-) \tag{63}$$

for scalar and vector x.

3.5. Predicting S

One variant utilizes the prognostic variables $(h, \mathbf{u}, S)$, and diagnoses $s \in \mathbb{W}_2$ from S and h.

3.5.1. Hamiltonian and Functional Derivatives

The Hamiltonian $\mathcal{H}[h, \mathbf{u}, S]$ is given as

$$\mathcal{H}[h, \mathbf{u}, S] = \frac{1}{2} \langle S, h + 2b \rangle + \frac{1}{2} \langle h \, \mathbf{u}, \mathbf{u} \rangle \tag{64}$$

Therefore, the auxiliary equations for the functional derivatives are:

$$\left\langle \hat{h}, \frac{\delta \mathcal{H}}{\delta h} \right\rangle = \left\langle \hat{h}, B \right\rangle = \left\langle \hat{h}, \frac{S}{2} + \frac{\mathbf{u} \cdot \mathbf{u}}{2} \right\rangle \tag{65}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\delta \mathcal{H}}{\delta \mathbf{u}} \right\rangle = \langle \hat{\mathbf{u}}, \mathbf{F} \rangle = \langle \hat{\mathbf{u}}, h \, \mathbf{u} \rangle \tag{66}$$

$$\left\langle \hat{S}, \frac{\delta \mathcal{H}}{\delta S} \right\rangle = \left\langle \hat{S}, T \right\rangle = \left\langle \hat{S}, \frac{h}{2} + b \right\rangle \tag{67}$$

Note that the functional derivatives live in the same space as their associated variable: $B, T \in \mathbb{W}_2$ and $\mathbf{F} \in \mathbb{W}_1$.

3.5.2. Discrete Brackets

Based on ideas from [66, 11], the discrete brackets are given as

$$\{\mathcal{A},\mathcal{B}\}_R = -\left\langle \frac{\delta\,\mathcal{A}}{\delta h}, \nabla\cdot\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}}\right\rangle + \left\langle \frac{\delta\,\mathcal{B}}{\delta h}, \nabla\cdot\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}}\right\rangle \tag{68}$$

$$\{\mathcal{A},\mathcal{B}\}_Q = -\left\langle \frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}}, q\frac{\delta\,\mathcal{B}^T}{\delta\,\mathbf{u}}\right\rangle \tag{69}$$

$$\{\mathcal{A},\mathcal{B}\}_S = \left\langle \nabla_H\frac{\delta\,\mathcal{A}}{\delta S}, s\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}}\right\rangle - \left\langle \nabla_H\frac{\delta\,\mathcal{B}}{\delta S}, s\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}}\right\rangle \tag{70}$$

$$+\left\langle [\frac{\delta\,\mathcal{B}}{\delta S}\frac{\delta\,\mathcal{A}}{\delta\,\mathbf{u}},\hat{\mathbf{n}}],\{s\}+c_f[s]\right\rangle_\Gamma - \left\langle [\frac{\delta\,\mathcal{A}}{\delta S}\frac{\delta\,\mathcal{B}}{\delta\,\mathbf{u}},\hat{\mathbf{n}}],\{s\}+c_f[s]\right\rangle_\Gamma \tag{71}$$

The diagnostic quantities q and s used in the brackets are obtained from:

$$\langle \hat{q}, hq\rangle = -\left\langle \nabla^T\hat{q},\mathbf{u}\right\rangle + \langle \hat{q}, f\rangle \tag{72}$$

$$\langle \hat{s}, hs\rangle = \langle \hat{s}, S\rangle \tag{73}$$

The stabilization term c_f is given by

$$c_f = \frac{\alpha}{2} \quad \text{if } \mathbf{F}\cdot\hat{\mathbf{n}} > 0 \qquad c_f = -\frac{\alpha}{2} \quad \text{if } \mathbf{F}\cdot\hat{\mathbf{n}} < 0 \tag{74}$$

where α is a parameter. When $\alpha = 0$, this results in a centered flux, while $\alpha = 1$ gives an upwind flux. These brackets are only anti-symmetric, they do not satisfy the Jacobi identity. However, anti-symmetry is sufficient to ensure conservation of $\mathcal{H}$.

3.5.3. Discrete Equations of Motion

The discrete equations of motion are then obtained through the discrete brackets by letting the functional $\mathcal{F}$ be simply the relevant variable multiplied by a test function from the appropriate space, exactly as in the continuous case. For example, the h evolution equation is obtained by setting

$$\mathcal{F} = \left\langle \hat{h}, h\right\rangle \tag{75}$$

This works since the test functions are, by definition, independent of time and therefore

$$\frac{d\,\mathcal{F}}{dt} = \left\langle \hat{h}, \frac{\partial h}{\partial t}\right\rangle \tag{76}$$

which is precisely the correct time derivative term. Applying this approach and using functional derivatives (65) - (67) in the Poisson brackets (68) - (70) yields:

$$\left\langle \hat{h}, \frac{\partial h}{\partial t}\right\rangle + \left\langle \hat{h}, \nabla\cdot\mathbf{F}\right\rangle = 0 \tag{77}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\partial\,\mathbf{u}}{\partial t}\right\rangle + \left\langle \hat{\mathbf{u}}, q\,\mathbf{F}^T\right\rangle - \langle \nabla\cdot\hat{\mathbf{u}}, B\rangle + \langle s\,\hat{\mathbf{u}}, \nabla_H T\rangle - \langle [T\,\hat{\mathbf{u}},\hat{\mathbf{n}}],\{s\}+c_f[s]\rangle_\Gamma = 0 \tag{78}$$

$$\left\langle \hat{S}, \frac{\partial S}{\partial t}\right\rangle - \left\langle \nabla_H\hat{S}, s\,\mathbf{F}\right\rangle + \left\langle [\hat{S}\,\mathbf{F},\hat{\mathbf{n}}],\{s\}+c_f[s]\right\rangle_\Gamma = 0 \tag{79}$$

Due to the choice of spaces, (77) holds pointwise, not just in an integral sense. Additionally, (65) can be directly substituted into (78), leaving only the auxiliary equations (66) and (67) to be solved. So the final system consists of the prognostic equations (77) - (79) and the diagnostic equations (66) - (67) and (72) - (73), with (65) substituted into (78). A discretization of the shallow water equations can be obtained from the thermal shallow water equations by setting $S = gh$ (i.e. $s = g$), just as in the continuous case. This eliminates the bracket $\{\mathcal{A}, \mathcal{B}\}_S$ and changes the Hamiltonian $\mathcal{H}$ and its associated functional derivatives. Therefore, the evolution equation for S and the diagnostic equations for s and T are eliminated. In fact, the spatial discretization scheme in [10] will be recovered.

3.5.4. Discretization of Linearized Equations

Again following the procedure in [50], and using the same reference state, the linearized equations can be obtained from the nonlinear equations. They are

$$\left\langle \hat{h}, \frac{\partial h}{\partial t} \right\rangle + H \left\langle \hat{h}, \nabla \cdot \mathbf{u} \right\rangle = 0 \tag{80}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \left\langle \hat{\mathbf{u}}, f\, \mathbf{u}^T \right\rangle - \frac{1}{2} \left\langle \nabla \cdot \hat{\mathbf{u}}, S \right\rangle - g \left\langle \nabla \cdot \hat{\mathbf{u}}, \frac{h}{2} \right\rangle = 0 \tag{81}$$

$$\left\langle \hat{S}, \frac{\partial S}{\partial t} \right\rangle + gH \left\langle \hat{S}, \nabla \cdot \mathbf{u} \right\rangle = 0 \tag{82}$$

This discretization of the linear equations is useful in the construction of a semi-implicit variant of the energy-conserving integrator discussed below. As for the nonlinear equations, when $S = gh$, the linear thermal shallow water equations reduce to the linear shallow water equations.

3.6. Predicting s

It is also possible to predict s instead of S, and we can place s in either $\mathbb{W}_2$ or $\mathbb{W}_0$. For both variants, the new Hamiltonian $\mathcal{H}'[h, \mathbf{u}, s]$ is

$$\mathcal{H}'[h, \mathbf{u}, s] = \frac{1}{2} \langle hs, h + 2b \rangle + \frac{1}{2} \langle h\, \mathbf{u}, \mathbf{u} \rangle \tag{83}$$

with functional derivatives

$$\left\langle \hat{h}, \frac{\delta\, \mathcal{H}'}{\delta h} \right\rangle = \left\langle \hat{h}, B' \right\rangle = \left\langle \hat{h}, sh + sb + \frac{\mathbf{u} \cdot \mathbf{u}}{2} \right\rangle \tag{84}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\delta\, \mathcal{H}'}{\delta\, \mathbf{u}} \right\rangle = \langle \hat{\mathbf{u}}, \mathbf{F} \rangle = \langle \hat{\mathbf{u}}, h\, \mathbf{u} \rangle \tag{85}$$

$$\left\langle \hat{s}, \frac{\delta\, \mathcal{H}'}{\delta s} \right\rangle = \langle \hat{s}, T' \rangle = \left\langle \hat{s}, \frac{h^2}{2} + hb \right\rangle \tag{86}$$

Equation (84) can be directly substituted into (92) or (97), but (85) and (86) must be solved. The $\{\mathcal{A}', \mathcal{B}'\}_R$ and $\{\mathcal{A}', \mathcal{B}'\}_Q$ brackets are given by

$$\{\mathcal{A}', \mathcal{B}'\}_R = -\left\langle \frac{\delta \mathcal{A}'}{\delta h}, \nabla \cdot \frac{\delta \mathcal{B}'}{\delta \mathbf{u}} \right\rangle + \left\langle \frac{\delta \mathcal{B}'}{\delta h}, \nabla \cdot \frac{\delta \mathcal{A}'}{\delta \mathbf{u}} \right\rangle \tag{87}$$

$$\{\mathcal{A}', \mathcal{B}'\}_Q = -\left\langle \frac{\delta \mathcal{A}'}{\delta \mathbf{u}}, q \frac{\delta \mathcal{B}'^T}{\delta \mathbf{u}} \right\rangle \tag{88}$$

which have the same form as (68) - (69) from the variant predicting S, while the $\{\mathcal{A}', \mathcal{B}'\}_s$ bracket will differ depending on the choice of space for s. The two variants are described below.

3.6.1. $s \in \mathbb{W}_0$

One variant places $s \in \mathbb{W}_0$. The discrete $\{\mathcal{A}', \mathcal{B}'\}_s$ bracket is

$$\{\mathcal{A}', \mathcal{B}'\}_s = \left\langle \frac{\nabla s}{h'}, \frac{\delta \mathcal{A}'}{\delta \mathbf{u}} \frac{\delta \mathcal{B}'}{\delta s} \right\rangle - \left\langle \frac{\nabla s}{h'}, \frac{\delta \mathcal{B}'}{\delta \mathbf{u}} \frac{\delta \mathcal{A}'}{\delta s} \right\rangle \tag{89}$$

where $h' \in \mathbb{W}_0$ is defined through

$$\langle \hat{s}, h' \rangle = \langle \hat{s}, h \rangle \tag{90}$$

This bracket is anti-symmetric, so energy will be conserved. Note that $T' \in \mathbb{W}_0$ since $s \in \mathbb{W}_0$. The equations of motion are:

$$\left\langle \hat{h}, \frac{\partial h}{\partial t} \right\rangle + \left\langle \hat{h}, \nabla \cdot \mathbf{F} \right\rangle = 0 \tag{91}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \left\langle \hat{\mathbf{u}}, q \mathbf{F}^T \right\rangle - \langle \nabla \cdot \hat{\mathbf{u}}, B' \rangle - \left\langle \frac{\nabla s}{h'}, \hat{\mathbf{u}} T' \right\rangle = 0 \tag{92}$$

$$\left\langle \hat{s}, \frac{\partial s}{\partial t} \right\rangle + \left\langle \frac{\nabla s}{h'}, \mathbf{F} \hat{s} \right\rangle = 0 \tag{93}$$

Since (77) holds pointwise, an evolution equation for h' is given as

$$\left\langle \hat{s}, \frac{\partial h'}{\partial t} \right\rangle + \langle \hat{s}, \nabla \cdot \mathbf{F} \rangle = 0 \tag{94}$$

This form is well-suited to the application of Streamline Upwind Petrov-Galerkin for the s equation, since there are no derivatives on $\hat{s}$. However, it is not clear how to keep energy conservation when using SUPG in this way, since the test function for the time derivative term would also be modified.

3.6.2. $s \in \mathbb{W}_2$

Alternatively, we can place $s \in \mathbb{W}_2$ and use ideas from [66]. Now $T' \in \mathbb{W}_2$ and the discrete $\{\mathcal{A}', \mathcal{B}'\}_s$ bracket is

$$\{\mathcal{A}', \mathcal{B}'\}_s = \left\langle \nabla \cdot (\frac{1}{h}\frac{\delta \mathcal{B}'}{\delta \mathbf{u}}\frac{\delta \mathcal{A}'}{\delta s}) - \nabla \cdot (\frac{1}{h}\frac{\delta \mathcal{A}'}{\delta \mathbf{u}}\frac{\delta \mathcal{B}'}{\delta s}), s \right\rangle - \left\langle [\frac{1}{h}\frac{\delta \mathcal{B}'}{\delta \mathbf{u}}\frac{\delta \mathcal{A}'}{\delta s}, \hat{\mathbf{n}}] - [\frac{1}{h}\frac{\delta \mathcal{A}'}{\delta \mathbf{u}}\frac{\delta \mathcal{B}'}{\delta s}, \hat{\mathbf{n}}], \{s\} + c_f[s] \right\rangle_\Gamma \tag{95}$$

This yields the equations of motion as

$$\left\langle \hat{h}, \frac{\partial h}{\partial t} \right\rangle + \left\langle \hat{h}, \nabla \cdot \mathbf{F} \right\rangle = 0 \tag{96}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \cdots + \left\langle \nabla \cdot (\frac{1}{h}\hat{\mathbf{u}}\, T'), s \right\rangle - \left\langle [\frac{1}{h}\hat{\mathbf{u}}\, T', \hat{\mathbf{n}}], \{s\} + c_f[s] \right\rangle_\Gamma = 0 \tag{97}$$

$$\left\langle \hat{s}, \frac{\partial s}{\partial t} \right\rangle - \left\langle \nabla \cdot (\frac{1}{h}\mathbf{F}\,\hat{s}), s \right\rangle + \left\langle [\frac{1}{h}\mathbf{F}\,\hat{s}, \hat{\mathbf{n}}], \{s\} + c_f[s] \right\rangle_\Gamma = 0 \tag{98}$$

3.6.3. Linearized Equations

For both approaches that predict s instead of S (with $s \in \mathbb{W}_2$ or $s \in \mathbb{W}_0$), the discretization of the linearized equations becomes

$$\left\langle \hat{h}, \frac{\partial h}{\partial t} \right\rangle + H \left\langle \hat{h}, \nabla \cdot \mathbf{u} \right\rangle = 0 \tag{99}$$

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \left\langle \hat{\mathbf{u}}, f\,\mathbf{u}^T \right\rangle - \left\langle \nabla \cdot \hat{\mathbf{u}}, gh + \frac{sH}{2} \right\rangle = 0 \tag{100}$$

$$\left\langle \hat{s}, \frac{\partial s}{\partial t} \right\rangle = 0 \tag{101}$$

3.7. Alternative form of $\{\mathcal{A}, \mathcal{B}\}_Q$

Following ideas from [11, 66, 67, 68], an alternative to diagnosing q and computing the nonlinear PV flux term in the $\frac{\partial \mathbf{u}}{\partial t}$ equation as

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \left\langle \hat{\mathbf{u}}, q\,\mathbf{F}^T \right\rangle + \cdots = 0 \tag{102}$$

is to directly discretize it as

$$\left\langle \hat{\mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle - \left\langle \nabla_H^T \left(\hat{\mathbf{u}} \cdot \frac{\mathbf{F}^T}{h} \right), \mathbf{u} \right\rangle + \left\langle \hat{\mathbf{u}}, f\frac{\mathbf{F}^T}{h} \right\rangle + \left\langle [\hat{\mathbf{u}} \cdot \frac{\mathbf{F}^T}{h}, \hat{\mathbf{n}}]^T, \{\mathbf{u}\} + c_f[\mathbf{u}] \right\rangle_\Gamma + \cdots = 0 \tag{103}$$

where ∇_H^T is the broken skew-gradient local to an element. This comes from a discrete $\{\mathcal{A}, \mathcal{B}\}_Q = \{\mathcal{A}', \mathcal{B}'\}_Q$ bracket given by

$$\{\mathcal{A}, \mathcal{B}\}_Q = \{\mathcal{A}', \mathcal{B}'\}_Q = \left\langle \nabla_H^T \left(\frac{1}{h}\frac{\delta \mathcal{A}}{\delta \mathbf{u}} \cdot (\frac{\delta \mathcal{B}}{\delta \mathbf{u}})^T \right), \mathbf{u} \right\rangle - \left\langle \frac{f}{h}\frac{\delta \mathcal{A}}{\delta \mathbf{u}}, (\frac{\delta \mathcal{B}}{\delta \mathbf{u}})^T \right\rangle - \left\langle [\frac{1}{h}\frac{\delta \mathcal{A}}{\delta \mathbf{u}} \cdot (\frac{\delta \mathcal{B}}{\delta \mathbf{u}})^T, \hat{\mathbf{n}}]^T, \{\mathbf{u}\} + c_f[\mathbf{u}] \right\rangle \tag{104}$$

This bracket is still anti-symmetric, so the discrete system will conserve energy $\mathcal{H}$. This approach is of interest because it offers an alternative to the q based variant that retains almost all of its properties, but might have different behavior in terms of the Hollingsworth instability.

3.8. Compatible Advection of s

Compatible advection of s requires that an initially uniform s remains so for all time.

3.8.1. S predicted

When S is predicted, compatible advection is equivalent to the requirement that the S equation reduces to the h equation when $s = 1$. Start by letting $s = 1$ in (79) to get

$$\left\langle \hat{S}, \frac{\partial S}{\partial t} \right\rangle - \left\langle \nabla_H \hat{S}, \mathbf{F} \right\rangle + \left\langle [\hat{S}\,\mathbf{F}, \hat{\mathbf{n}}], 1 \right\rangle_\Gamma = 0 \tag{105}$$

where $[s] = 0$ and $\{s\} = 1$ have been used. The last two terms can be combined using integration by parts (since $\mathbf{F}$ has continuous normals) to yield

$$\left\langle \hat{S}, \frac{\partial S}{\partial t} \right\rangle + \left\langle \hat{S}, \nabla \cdot \mathbf{F} \right\rangle = 0 \tag{106}$$

Finally, the time derivative of (73) gives

$$\left\langle \hat{s}, \frac{\partial (hs)}{\partial t} \right\rangle = \left\langle \hat{s}, \frac{\partial S}{\partial t} \right\rangle \tag{107}$$

Equations (106) and (107) can be combined (using $s = 1$ and $\hat{s} = \hat{h}$, the latter because they are in the same space) to yield

$$\left\langle \hat{S}, \frac{\partial h}{\partial t} \right\rangle + \left\langle \hat{S}, \nabla \cdot \mathbf{F} \right\rangle = 0 \tag{108}$$

This is equivalent to (77) since $\hat{S}$ and $\hat{h}$ are in the same space.

3.8.2. s predicted

Here we consider what happens to the s equation when s is a constant. We start by noting that $\nabla s = \nabla_H s = [s] = 0$. Simple inspection of (93) gives

$$\left\langle \hat{s}, \frac{\partial s}{\partial t} \right\rangle = 0 \tag{109}$$

For (98), similar calculations to those for the S variant, involving integration by parts also give $\left\langle \hat{s}, \frac{\partial s}{\partial t} \right\rangle = 0$. Therefore, an initially constant s will remain so for all time.

3.9. Implied Vorticity Equation and PV Compatibility

If we let $\hat{\mathbf{u}} = -\nabla^T \hat{q}$ and time differentiate (72), a discrete PV equation emerges by combining this with (78):

$$\left\langle \hat{q}, \frac{\partial (hq)}{\partial t} \right\rangle + \langle \hat{q}, \nabla \cdot (q\,\mathbf{F}) \rangle - \left\langle s \nabla^T q, \nabla_H T \right\rangle + \left\langle [T \nabla^T q, \hat{\mathbf{n}}], \{s\} + c_f[s] \right\rangle_\Gamma = 0 \tag{110}$$

As expected, the term involving B has dropped out, and when s is a constant the terms involving T will also be zero. Showing the latter again requires integration by parts and exploiting the fact that $\nabla^T q \in \mathbb{W}_1$ and therefore has continuous normals. This is a demonstration that the scheme does not have spurious sources of vorticity. Similarly, using (92) yields

$$\left\langle \hat{q}, \frac{\partial (hq)}{\partial t} \right\rangle - \left\langle \nabla \hat{q}^T, q\,\mathbf{F}^T \right\rangle + \left\langle \frac{\nabla s}{h'}, T' \nabla^T \hat{q} \right\rangle = 0 \tag{111}$$

and using (97) yields

$$\left\langle \hat{q}, \frac{\partial (hq)}{\partial t} \right\rangle - \left\langle \nabla^T \hat{q}, q\,\mathbf{F}^T \right\rangle - \left\langle \nabla \cdot (\frac{1}{h} \nabla^T \hat{q} T'), s \right\rangle + \left\langle [\frac{1}{h} \nabla^T \hat{q} T', \hat{\mathbf{n}}], \{s\} + c_f[s] \right\rangle_{\Gamma_I} = 0 \tag{112}$$

which both have the same property that the terms involving T drop out when s is a constant. All of these equations reduce to

$$\left\langle \hat{q}, \frac{\partial (hq)}{\partial t} \right\rangle - \left\langle \nabla^T \hat{q}, q\,\mathbf{F}^T \right\rangle = 0 \tag{113}$$

when s is a constant, which is the same as in [10]. Therefore by letting $q = 1$ it is clear that the advection of potential vorticity is compatible with the continuity equation, since (77) holds pointwise. It is also easy to show that potential enstrophy $\mathcal{PE} = \frac{1}{2} \langle hq, q \rangle$ is conserved when s is a constant. The implied vorticity equation for the direct variant (103) is

$$\left\langle \hat{\eta}, \frac{\partial \eta}{\partial t} \right\rangle + \left\langle \nabla_H^T \left(\nabla^T \hat{\eta} \cdot \frac{\mathbf{F}^T}{h} \right), \mathbf{u} \right\rangle - \left\langle \nabla^T \hat{\eta}, f \frac{\mathbf{F}^T}{h} \right\rangle - \left\langle [\nabla^T \hat{\eta} \cdot \frac{\mathbf{F}^T}{h}, \hat{\mathbf{n}}]^T, \{\mathbf{u}\} + c_f[\mathbf{u}] \right\rangle_\Gamma + \cdots = 0 \tag{114}$$

which does not have PV compatibility or potential enstrophy conservation, since there is no longer an associated diagnostic q that appears in the implied vorticity equation. However, the resulting linear equations are the same (seen by letting $\mathbf{u} = 0$ and $h = H$ in (103)), so the discretization still supports steady geostrophic modes and does not produce spurious vorticity, and the discrete dispersion relationship does not change.

3.10. Conservation Properties and Casimirs

The discretizations presented above all conserve total energy $\mathcal{H}$ through the anti-symmetry of the discrete brackets. However, only specific choices of time integrators that are capable of handling polynomial Hamiltonians of cubic order will maintain this conservation. There are also three conserved Casimirs: total mass, total potential vorticity and total buoyancy. These Casimirs are at most quadratic.

3.10.1. Mass Conservation

For all discretizations the mass is given by

$$\mathcal{M} = \langle 1, h \rangle \tag{115}$$

Mass conservation is demonstrated by letting $\hat{h} = 1$ in (77). This is a flux-form conservation law and therefore any consistent time integrator will continue to conserve mass.

3.10.2. Total Potential Vorticity

The variant that diagnoses q and uses it in the $\mathbf{u}$ equation conserves a total potential vorticity (or in other words, an absolute vorticity) given by

$$\mathcal{PV} = \langle h, q \rangle \tag{116}$$

This is seen by letting $\hat{q} = 1$ in (110), (111) or (112). The variant in Section 3.7 instead conserves

$$\mathcal{PV} = \langle 1, \eta \rangle \tag{117}$$

where $\eta \in \mathbb{W}_0$ is defined by

$$\langle \hat{\eta}, \eta \rangle = -\left\langle \nabla^T \hat{\eta}, \mathbf{u} \right\rangle + \langle \hat{\eta}, f \rangle \tag{118}$$

This can be seen by letting $\hat{\eta} = 1$ in (114). Both quantities arise from flux-form conservation laws and any consistent time integrator will preserve them.

3.10.3. Total Buoyancy

The variant that predicts S will conserve a total buoyancy of the form

$$\mathcal{B} = \langle 1, S \rangle \tag{119}$$

demonstrated by setting $\hat{S} = 1$ in (79). Again, this is a flux-form conservation law and any consistent time integrator will preserve total buoyancy of this form. The variant that predicts $s \in \mathbb{W}_2$ will conserve

$$\mathcal{B} = \langle h, s \rangle \tag{120}$$

(set $\hat{s} = h$ and $\hat{h} = s$, and combine (77) and (93)). Finally, the variant that predicts $s \in \mathbb{W}_0$ will conserve

$$\mathcal{B} = \langle h', s \rangle \tag{121}$$

which is seen by combining (93) and (94) with $\hat{s} = h'$ and $\hat{h'} = s$. However, the last two variants require a time integrator that preserves quadratic Casimirs, since S is no longer being predicted directly using a flux-form scheme. Fortunately, the EC2 (and it's semi-implicit variant EC2-SI) integrators from Section 4 satisfy this. Unfortunately, standard explicit time integrators such as those of the Runge-Kutta or Adams-Bashford families do not.

4. Time Discretization

Under the assumption of continuity in time, the spatial discretization scheme developed in Section 3 obtains all of the desirable properties detailed in Section 3.1. However, the complete model also requires a time discretization scheme, and it is not clear that it is still possible to obtain these properties when the two are combined. Specifically, unless carefully designed, a time discretization scheme will fail to conserve quadratic and higher order invariants; and will lead to alterations in the discrete linear modes. Of the conserved quantities of interest, the higher-order ones are the Hamiltonian $\mathcal{H}$ (cubic); and for the variants predicting s, the total buoyancy Casimir $\mathcal{B}$ (quadratic). A time integrator and thus fully discrete scheme that conserves arbitrary $\mathcal{H}$ and quadratic Casimirs to machine precision is presented below. Work is currently ongoing on evaluating the fully discrete dispersion relationship and linear modes for this scheme.

4.1. Energy Conserving Time Integrator

We start by noting that the discrete equations of motion for all the variants can be written as:

$$\frac{\partial x}{\partial t} = \mathbb{J}\frac{\delta \mathcal{H}}{\delta x} \tag{122}$$

where x is a vector of the prognostic variables (basis coefficients), $\mathbb{J}$ is a skew-symmetric, singular matrix that depends on x and $\frac{\delta \mathcal{H}}{\delta x}$ are the functional derivatives of the discrete Hamiltonian $\mathcal{H}$. The matrix $\mathbb{J}$ arises from the discrete bracket as

$$\{\mathcal{A}, \mathcal{B}\} = \left\langle \frac{\delta \mathcal{A}}{\delta x}, \mathbb{J}\frac{\delta \mathcal{B}}{\delta x} \right\rangle \tag{123}$$

and therefore the nullspace of $\mathbb{J}$ is the discrete Casimirs. This is a specific instance of the general theorem proved in [69] that any system of ODE's possessing a conserved quantity can be written as (122). In the ODEs literature, these systems are known as Poisson (or gradient) systems. Fortunately, there exists a rich variety of energy conserving time integrators for Poisson systems. One such integrator (hereafter called EC2, from [70]; see also [42] for an example of its use for the shallow water equations) is the following:

$$\frac{x^{n+1} - x^n}{\Delta t} = \mathbb{J}\left(\frac{x^{n+1} + x^n}{2}\right) \int_0^1 \frac{\delta \mathcal{H}}{\delta x}(x^n + \tau(x^{n+1} - x^n))d\tau \tag{124}$$

which is a fully implicit, second-order accurate time integrator. As discussed in [70], this is an example of a continuous stage partitioned Runge-Kutta (CSPRK) method, and it is also an example of a discrete gradient or average method field method.

It is clear that $\mathcal{H}$ is conserved by anti-symmetry, and that quadratic Casimirs will also be conserved. The latter rests on the fact that for a quadratic Casimir, $\int_0^1 \frac{\delta \mathcal{C}}{\delta x}(x^\tau)d\tau$ can be evaluated with a single quadrature point at $\tau = 0.5$ and therefore both $\int_0^1 \frac{\delta \mathcal{F}}{\delta x}(x^\tau)d\tau$ and $\mathbb{J}$ are evaluated at x^*. Unfortunately, this does not generalize to higher-order Casimirs. This integrator can be viewed as a generalization of the implicit midpoint rule: when the

Hamiltonian $\mathcal{H}$ is quadratic, a single quadrature point at $\tau = 0.5$ is sufficient and the scheme reduces to the implicit midpoint rule. Practical implementation of this integrator for general $\mathcal{H}$ requires the evaluation of the integral on the right hand side of (124). When $\mathcal{H}$ is polynomial, as is the case for the thermal shallow water equations, this can be done exactly with a Gaussian quadrature rule of the appropriate degree. Proceed by discretizing the integral on the right-hand side of (124) using a p-pt Gaussian quadrature rule on $[0, 1]$ with weights γ and points τ_i, where $i = [1, \ldots, p]$. Define

$$\begin{aligned} \hat{\gamma}_i &= \gamma_i \Delta t && (125)\\ \delta x^{n+1} &= x^{n+1} - x^n && (126)\\ x^* &= x^n + \frac{\delta x^{n+1}}{2} && (127)\\ x^i &= x^{\tau_i} = x^n + \tau_i \delta x^{n+1} && (128) \end{aligned}$$

Then the time-discrete system of nonlinear equations is given by

$$\delta x^{n+1} = \mathbb{J}(x^*) \sum_i \hat{\gamma}^i \frac{\delta \mathcal{H}}{\delta x}(x^i) = \mathbb{J}(x^*) \overline{\frac{\delta \mathcal{H}}{\delta x}(x^i)} \tag{129}$$

where $\overline{\frac{\delta \mathcal{H}}{\delta x}(x^i)} = \sum_i \hat{\gamma}^i \frac{\delta \mathcal{H}}{\delta x}(x^i)$, along with any auxiliary equations needed for $\overline{\frac{\delta \mathcal{H}}{\delta x}(x^i)}$ and any diagnostic equations needed for $\mathbb{J}(x^*)$. Now, since $\mathcal{H}$ for the thermal shallow water equations is cubic, let $p = 2$. In fact, it is even possible to use different numbers of quadrature points for different parts of the Hamiltonian $\mathcal{H}$, although this suggestion is not pursued here. This is useful, for example, in the shallow water equations or the thermal shallow water equations predicting S, when the potential energy $\mathcal{H}_P$ is only quadratic and therefore can be exactly integrated with just one quadrature point. For simplicity, we show only the fully discrete system for the variant predicting s with $s \in \mathbb{W}_0$ and using the q form of PV flux. Here we also drop the primes on B and T to ease the notation, and let $m = h'$. The other variants give very similar fully discrete equations. A fully discrete model of prognostic equations (91) - (93), the functional derivative equations (84) -(86) and the auxiliary equations (72) and (94) using this integrator results in the following system:

$$\left\langle \hat{h}, \delta h^{n+1} \right\rangle + \left\langle \hat{h}, \nabla \cdot \overline{\mathbf{F}} \right\rangle = 0 \tag{130}$$

$$\left\langle \hat{\mathbf{u}}, \delta \mathbf{u}^{n+1} \right\rangle + \left\langle \hat{\mathbf{u}}, q^* \overline{\mathbf{F}}^T \right\rangle - \left\langle \nabla \cdot \hat{\mathbf{u}}, \overline{B} \right\rangle - \left\langle \frac{1}{m^*} \nabla s^*, \overline{T} \right\rangle = 0 \tag{131}$$

$$\left\langle \hat{s}, \delta s^{n+1} \right\rangle + \left\langle \hat{s} \frac{1}{m^*} \nabla s^*, \overline{\mathbf{F}} \right\rangle = 0 \tag{132}$$

$$\begin{aligned} \langle \hat{q}, h^* q^* \rangle &= -\left\langle \nabla^T \hat{q}, \mathbf{u}^* \right\rangle + \langle \hat{q}, f \rangle && (133)\\ \langle \hat{s}, m^* \rangle &= \langle \hat{s}, h^* \rangle && (134) \end{aligned}$$

$$\left\langle \hat{\mathbf{u}}, \overline{\mathbf{F}} \right\rangle = \left\langle \hat{\mathbf{u}}, \hat{\gamma}_1 h^1 \, \mathbf{u}^1 + \hat{\gamma}_2 h^2 \, \mathbf{u}^2 \right\rangle \tag{135}$$

$$\left\langle \hat{h}, \overline{B} \right\rangle = \left\langle \hat{h}, \hat{\gamma}_1 \left(s^1 h^1 + s^1 b + \frac{\mathbf{u}^1 \cdot \mathbf{u}^1}{2} \right) + \hat{\gamma}_2 \left(s^2 h^2 + s^2 b + \frac{\mathbf{u}^2 \cdot \mathbf{u}^2}{2} \right) \right\rangle \tag{136}$$

$$\left\langle \hat{s}, \overline{T} \right\rangle = \left\langle \hat{s}, \hat{\gamma}_1 \left(\frac{(h^1)^2}{2} + h^1 b \right) + \hat{\gamma}_2 \left(\frac{(h^2)^2}{2} + h^2 b \right) \right\rangle \tag{137}$$

As in the semi-discrete case, (136) can be directly substituted into (131), giving a nonlinear system $F(x) = 0$ in 7 variables: δh^{n+1}, $\delta \, \mathbf{u}^{n+1}$, δs^{n+1}, $\overline{\mathbf{F}}$, $\overline{T}$, m^*, q^*. This can be solved, for example, using a Newton-Krylov method.

4.2. Semi-Implicit (Quasi-Newton) Implementation

Although the fully implicit integrator gives the desired properties (such as conservation), it is computationally expensive, and the Jacobian matrix that appears in the Newton-Krylov method is difficult to precondition. Therefore, following [65], an alternative approach is taken (called hereafter EC2-SI) to reduce the computational cost while retaining the properties. Instead of a Newton method with the full Jacobian, a quasi-Newton (also called inexact Newton) approach using a simplified Jacobian is employed: the Jacobian from discretizing the linearized thermal shallow water equations (80) - (82) or (99) - (101) using an implicit midpoint time integrator. This has the advantage that the Jacobian becomes constant in time (and therefore must be calculated only once per simulation instead of at each outer solver iteration), and also that the resulting linear problem involves only 3 variables: δh^{n+1}, $\delta \, \mathbf{u}^{n+1}$, δs^{n+1} rather than the full set of 7 variables. This must be supplemented with auxiliary solves for q^*, m^*, $\overline{\mathbf{F}}$ and $\overline{T}$ using the latest guess for δx^{n+1}, done before evaluating the residual. The final linear system that results is much faster to solve and easier to precondition than the original system. In fact, since the δh^{n+1} equation holds pointwise, this problem could be further simplified by directly substituting it into the relevant other equations, but this is not pursued. Since the same residual is used, at convergence the solution is the same. The choice of simplified Jacobian will affect only the number of iterations required to reach convergence (and in some cases, whether convergence can be reached at all). This quasi-Newton approach is strongly related to the semi-implicit[2] time stepping schemes described in [65, 71, 72].

Specifically, for (130) - (137), the Jacobian from the implicit midpoint discretization of the linearized thermal shallow water equations (99) - (101) is used

$$\left\langle \hat{h}, \hat{\delta h} \right\rangle + \frac{H \Delta t}{2} \left\langle \hat{h}, \nabla \cdot \hat{\delta v} \right\rangle \tag{138}$$

$$\left\langle \hat{\mathbf{u}}, \hat{\delta v} \right\rangle + \frac{f \Delta t}{2} \left\langle \hat{\mathbf{u}}, (\hat{\delta v})^T \right\rangle - \frac{H \Delta t}{4} \left\langle \nabla \cdot \hat{\mathbf{u}}, \hat{\delta s} \right\rangle - \frac{g \Delta t}{2} \left\langle \nabla \cdot \hat{\mathbf{u}}, \hat{\delta h} \right\rangle \tag{139}$$

[2]The use of a simplified Jacobian to obtain a simpler (inner) linear system, usually one related to a linearization at the continuous level, is often referred to as a semi-implicit method in the atmospheric dynamical core literature. However, confusingly this term is also used to refer to single-step or diagonally implicit methods.

$$\left\langle \hat{s}, \hat{\delta s} \right\rangle \tag{140}$$

where $\hat{\delta x} = (\hat{\delta h}, \hat{\delta v}, \hat{\delta s})$ is a trial function for the mixed space $\mathbb{W}_2 \otimes \mathbb{W}_1 \otimes \mathbb{W}_0$, so that the above is a variational 2-form in $\hat{x}$ and $\hat{\delta x}$. Again, we show only the variant predicting s with $s \in \mathbb{W}_0$. Written in matrix form, the Jacobian $\mathbf{J}$ is

$$\mathbf{J} = \begin{bmatrix} \mathbf{M}_h & \frac{H\Delta t}{2}\mathbf{D} & 0 \\ \frac{-g\Delta t}{2}\mathbf{G}_h & \mathbf{M}_u + \frac{f\Delta t}{2}\mathbf{C} & \frac{-H\Delta t}{4}\mathbf{G}_s \\ 0 & 0 & \mathbf{M}_s \end{bmatrix} \tag{141}$$

where

$$\begin{aligned} \mathbf{M}_h &= \left\langle \hat{h}, \hat{\delta h} \right\rangle & \mathbf{M}_u &= \left\langle \hat{\mathbf{u}}, \hat{\delta v} \right\rangle & \mathbf{M}_s &= \left\langle \hat{s}, \hat{\delta s} \right\rangle \\ \mathbf{C} &= \left\langle \hat{\mathbf{u}}, (\hat{\delta v})^T \right\rangle & \mathbf{D} &= \left\langle \hat{h}, \nabla \cdot \hat{\delta v} \right\rangle & \mathbf{G}_h &= \left\langle \nabla \cdot \hat{\mathbf{u}}, \hat{\delta h} \right\rangle & \mathbf{G}_s &= \left\langle \nabla \cdot \hat{\mathbf{u}}, \hat{\delta s} \right\rangle \end{aligned} \tag{142}$$

This is a suboptimal simplified Jacobian (see Section 5.1), but it suffices to demonstrate the ability of the scheme to conserve quantities to machine-precision even with a simplified Jacobian.

5. Results

5.1. Implementation

The various discretizations given above were implemented using the Themis software framework ([73]), which shares a front-end with the Firedrake [74] project consisting of UFL (Unified Form Language, [75]), TSFC (Two-Stage Form Compiler, [76, 77]), FInAT (FInAT/FInAT: a smarter library of finite elements, [78]), COFFEE (COmpiler For Fast Expression Evaluation,[79, 80]) and FIAT (FInite element Automatic Tabulator, [81]). In fact, as currently implemented the model code runs using both Themis and Firedrake without changes. The resulting numerical kernels are used along with the Portable Extensible Toolkit for Scientific Computation (PETSc, [82]) to solve the various linear and nonlinear systems that arise in the fully discrete schemes. Themis is a software framework designed for parallel, high-performance automated discretization of variational forms using tensor-product Galkerin methods on multipatch structured-grids; with an emphasis on compatible Galerkin methods. Currently support exists for the MGD_n and $Q_r^-\Lambda^k$ families on single patch grids, with extensions to multipatch grids and additional compatible Galerkin families under development. All of the results in this paper use the MGD_3 family and the EC2-SI time integrator with 2-pt Gaussian quadrature in the time integrator and 5-pt Gaussian quadrature in the spatial integrals, and were run on a workstation using 9 MPI threads. The nonlinear system arising from the quasi-Newton EC2-SI time integrator was solved using a line search Newton method, with all associated linear systems solved using either the conjugate gradient (CG, for symmetric matrices) or generalized gradient minimal residual method (GMRES,

for non-symmetric matrices) with Jacobi preconditioning. This included the block matrix system occurring in the inner solve for the quasi-Newton integrator, which was treated in a monolithic manner. The use of Jacobi preconditioning is clearly suboptimal, especially for the block-matrix system. We intend to explore more sophisticated preconditioning options that exploit both the block structure and the MGD_n basis functions in future work, and to study the effects of grid resolution and order of accuracy on the number of inner iterations required. For the double vortex and thermal instability test cases we also performed runs for all six variants using the full Jacobian rather than the simplified Jacobian. When using the full Jacobian the Newton method took approximately 4-5 iterations to each convergence, while quasi-Newton method using the simplified Jacobian took approximately 30-40 iterations. This is a demonstration of the suboptimality of the simplified Jacobian, and we intend to explore more accurate simplified Jacobians in future work.

Three separate test cases were run to demonstrate conservation properties, correct solution behavior and convergence of the proposed schemes. Differences between the various schemes presented in Section 3 will be highlighted. The first test case, described in Section 5.2, is an analogue of the Williamson Test Case 2 [83] on the f-plane for the thermal shallow water equations, and is intended mainly to show convergence and the ability of the schemes to maintain a thermogeostrophically balanced state. The second, in Section 5.3, is a pair of vortices in a zonally varying buoyancy field, and is used to investigate conservation properties for complicated, non-linear flow; and also compare the various discretization options. The final, in Section 5.4, is based on thermogeostrophic instability from [6]. It is used to look at conservation properties and the ability to correctly simulate the onset of the flow instability identified in [6]. Table 5.1 details the various test cases and settings used. For Section 5.2 and 5.4, the initial condition is in thermogeostrophic balance (see [6] and Appendix Appendix A), with Section 5.4 adding a perturbation to start the flow. For all tests, the time step was chosen as $\Delta t = \frac{C\Delta x}{\sqrt{gH_0}}$ where $\Delta x = \frac{L}{nx}$ and $C = 2.0$ (it is based on the gravity wave CFL condition). The parameters L, g and H_0 are defined separately for each test case below, and were chosen to correspond with those used in [84]. All test cases were run on a uniform square mesh of nx by ny quadrilateral elements. It is worth noting that all simulations were performed without any numerical dissipation.

5.2. Zonal Thermogeostrophic Balance

This test is designed to test the ability of the schemes to maintain a state of thermogeostrophic balance, and to assess the convergence of the schemes (since there is an analytic solution available). The initial condition is a zonally symmetric, thermogeostrophically balanced zonal flow without topography ($b = 0$):

$$h = H_0 - \frac{afu_0}{g}\sin(\frac{y}{a}) \qquad u = u_0\cos(\frac{y}{a}) \qquad v = 0 \qquad s = g\left(1 + \frac{cH_0^2}{h^2}\right) \tag{143}$$

This state was obtained following the procedure in Appendix A (let h and $\mathbf{u}$ be in geostrophic balance, and then find s such that thermogeostrophic balance holds). When $c = 0$, $s = g$ and a geostrophically balanced zonal flow is recovered. The domain $\Omega = [0, L]^2$ is doubly periodic,

Test Case	nx	ny	Δx	N	Δt	α_s	α_q
Zonal Thermogeostrophic Balance (5.2)	10	10	4003km	340	16560s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	15	15	2670km	500	11040s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	30	30	1334km	1000	5520s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	45	45	890km	1500	3680s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	60	60	667km	2000	2760s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	90	90	444km	3000	1840s	0.0	0.0
Zonal Thermogeostrophic Balance (5.2)	120	120	333km	4000	1380s	0.0	0.0
Double Vortex (5.3)	120	120	42km	500	486s	0.0	0.0
Thermal Instability (5.4)	120	120	0.0666	500	0.0666	0.0	0.0

Table 4: Table of parameters for the test cases, where nx and ny are the number of elements in the x and y directions, Δx is the width of each element, N is the number of time steps, Δt is the length of a time step and α_s/α_q are the upwinding parameters for the $\{\mathcal{A},\mathcal{B}\}_S$ or $\{\mathcal{A}',\mathcal{B}'\}_s$ and $\{\mathcal{A},\mathcal{B}\}_Q = \{\mathcal{A}',\mathcal{B}'\}_Q$ brackets, respectively. Note that the thermal instability test case was non-dimensionalized.

with $L = 2\pi a$. The parameters, chosen to correspond with [48], are $f = 0.00006147s^{-1}$, $a = 6371120$m, $H_0 = 5960$m, $g = 9.80616ms^{-2}$, $u_0 = 20ms^{-1}$ and $c = .05$.

This test was run with for a range of grid sizes and number of time steps as shown in Table 5.1. These were chosen so that the same total simulation time was performed for all choices. The L_2 norms of the difference between the initial $(h, \mathbf{u}, s)$ and the final $(h, \mathbf{u}, s)$ are found in Figure 3. From the figures, it is clear that all of the variants are converging at between fourth and fifth order. This is significantly better than expected, and likely a result of superconvergence due to the uniform grid and alignment of the test case with the grid lines. The flattening at the end of convergence covers indicates that the limits of numerical precision have been reached.

5.3. Double Vortex Test Case

This test case is based on the double vortex test case from [84]. The domain $\Omega = [-\frac{L}{2}, \frac{L}{2}]^2$ is doubly periodic without topography ($b = 0$). The initial conditions are given by

$$h = H_0 - \Delta h \left[e^{-0.5((x_1')^2+(y_1')^2)} + e^{-0.5((x_2')^2+(y_2')^2)} - \frac{4\pi\sigma_x\sigma_y}{L^2} \right] \tag{144}$$

$$u = \frac{-g\Delta h}{f\sigma_y} \left[y_1'' e^{-0.5((x_1')^2+(y_1')^2)} + y_2'' e^{-0.5((x_2')^2+(y_2')^2)} \right] \tag{145}$$

$$v = \frac{g\Delta h}{f\sigma_x} \left[x_1'' e^{-0.5((x_1')^2+(y_1')^2)} + x_2'' e^{-0.5((x_2')^2+(y_2')^2)} \right] \tag{146}$$

$$s = g\left(1 + 0.05 \sin\left[\frac{2\pi}{L}(x - xc)\right]\right) \tag{147}$$

where $xc = \frac{L}{2}$ and

$$x_1' = \frac{L}{\pi\sigma_x}\sin\left[\frac{\pi}{L}(x - xc_1)\right] \qquad x_2' = \frac{L}{\pi\sigma_x}\sin\left[\frac{\pi}{L}(x - xc_2)\right] \tag{148}$$

$$y_1' = \frac{L}{\pi\sigma_y}\sin\left[\frac{\pi}{L}(y - yc_1)\right] \qquad y_2' = \frac{L}{\pi\sigma_y}\sin\left[\frac{\pi}{L}(y - yc_2)\right] \tag{149}$$

$$x_1'' = \frac{L}{2\pi\sigma_x}\sin\left[\frac{2\pi}{L}(x - xc_1)\right] \qquad x_2'' = \frac{L}{2\pi\sigma_x}\sin\left[\frac{2\pi}{L}(x - xc_2)\right] \tag{150}$$

$$y_1'' = \frac{L}{2\pi\sigma_y}\sin\left[\frac{2\pi}{L}(y - yc_1)\right] \qquad y_2'' = \frac{L}{2\pi\sigma_y}\sin\left[\frac{2\pi}{L}(y - yc_2)\right] \tag{151}$$

This is not in thermogeostrophic or geostrophic balance. The centers of the two vortices are given by

$$xc_1 = (0.5 - ox)L \qquad xc_2 = (0.5 + ox)L \tag{152}$$

$$yc_1 = (0.5 - oy)L \qquad yc_2 = (0.5 + oy)L \tag{153}$$

The parameters are $L = 5000$km, $f = 0.00006147s^{-1}$, $H_0 = 750$m, $\Delta h = 75$m, $g = 9.80616ms^{-2}$, $\sigma_x = \sigma_y = \frac{3}{40}L$ and $ox = oy = 0.1$. This test was run for the grid size and time step detailed in Table 5.1. Plots of the conservation properties are found in Figure 5, demonstrating that all of the variants are conserving $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ to machine precision. Here $\mathcal{H}_A$ is the available energy, rather than the total energy, since available energy is the dynamically active part. The available energy is defined as

$$\mathcal{H}_A = \mathcal{H} - \mathcal{H}_{UA} = \mathcal{H} - \frac{1}{2}\langle h_0, S_0\rangle \tag{154}$$

where $h_0 = \frac{\mathcal{M}}{Lx*Ly}$ and $S_0 = \frac{\mathcal{B}}{Lx*Ly}$. The second term on the right hand side is the unavailable energy $\mathcal{H}_{UA}$ associated with a state of rest and constant h and s. The initial potential vorticity q and the buoyancy s are found in Figure 4, while the corresponding quantities at $N = 250$ are given in Figures 6 and 7. This test case is a direct cascade without diffusion. As the simulation proceeds, progressively smaller scales are generated, and even by $N = 250$ there is significant development of small scale features. The ability of the schemes to successfully complete these runs without any numerical viscosity is a demonstration of their robustness. For this test case, there does not appear to be any significant difference between the variants.

5.4. Thermal Instability

The final test case is thermal instability, from [6]. To define this test case, it is useful to non-dimensionalize the equations with a length scale L, a velocity scale U, a fluid height scale H_0 and a buoyancy scale s_0. From these, a Rossby number $Ro = \frac{U}{fL}$ and a Burgers number $Bu = \frac{s_0 H_0}{f^2 L^2}$ are naturally defined. The variables are then scaled as

$$\frac{h}{H_0} = 1 + \frac{Ro}{Bu}H \qquad \frac{s}{s_0} = 1 + 2\frac{Ro}{Bu}B \qquad \frac{v}{U} = 1 \tag{155}$$

Now consider an axisymmetric state in thermogeostrophic balance, such that the non-dimensional variables are functions only of radial distance r, not of polar angle ψ; and we assume only an azimuthal velocity $v_a(r) = UV(r)$. The non-dimensional variables $H(r)$, $B(r)$ and $V(r)$ are given by

$$\begin{aligned} H(r) &= 0 && (156)\\ B(r) &= -\int_r^\infty (1 + \frac{RoV}{r'})V dr' && (157)\\ V(r) &= re^{\frac{1-r^\beta}{\beta}} && (158) \end{aligned}$$

This corresponds to choosing $\alpha = 0$ in [85]. The velocity vector is then given by

$$\mathbf{u} = v_a\hat{\boldsymbol{\psi}} = v_a(-\sin\psi, \cos\psi) \tag{159}$$

The domain $\Omega = [-\frac{D}{2}, \frac{D}{2}]^2$ is doubly periodic, with no topography. The parameters are $D = 4$, $g = H_0 = f = s_0 = L = 1.0$, $U = 0.1$ and $\beta = 2$. This gives $Ro = 0.1$ and $Bu = 1.0$. For $\beta = 2$ the equation for $B(r)$ can be solved explicitly to yield

$$B(r) = -\left[e^{\frac{1-r^2}{2}} + \frac{Ro}{2}e^{1-r^2}\right] \tag{160}$$

Re-dimensionalizing, the variables are given by

$$\begin{aligned} h &= H_0 && (161)\\ u &= -Ure^{\frac{1-r^\beta}{\beta}}\sin\psi && (162)\\ v &= Ure^{\frac{1-r^\beta}{\beta}}\cos\psi && (163)\\ s &= s_0 - 2\frac{s_0 Ro}{Bu}\left[e^{\frac{1-r^2}{2}} + \frac{Ro}{2}e^{1-r^2}\right] && (164) \end{aligned}$$

This is different to [6], where the choice $\beta = 3$ was made instead. We chose $\beta = 2$ instead, to facilitate the explicit solution of the $B(r)$ equation. To this thermogeostrophically balanced state a perturbation of the form

$$\begin{aligned} dh &= 0.01 s_f \cos(l\phi) && (165)\\ ds &= -0.01 s_f \cos(l\phi) && (166)\\ du &= -0.01 s_f \cos(l\phi) && (167)\\ dv &= -0.01 s_f \cos(l\phi) && (168) \end{aligned}$$

was added, where

$$s_f = -e^{-60(r-rc)^2}\sin(6\pi(r - rc)) \tag{169}$$

with $l = 4$ and $rc = 0.5$. Plots of the initial buoyancy s and the perturbation are found in Figure 8. The perturbation is an approximation to the most unstable linear mode structures

discussed in [6]. This test was run for the grid size and time step detailed in Table 5.1. The buoyancy at $N = 180$ is given in Figure 9. Similar to [6], there is an initial period of growth reflecting the wavenumber of the perturbation ($l = 4$ in this case). As simulation progresses, nonlinear saturation of the instability occurs (not shown). Despite the appearance of extremely small scales and nonlinear saturation, the schemes are robust and stable. Again, there appears to be little difference between the six variants of the scheme. The conservation properties for $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ are found in Figure 10. Just like the double vortex test case, there is conservation to machine precision for all these invariants.

6. Conclusions and Future Work

This work represents the beginning of the development of a structure-preserving atmospheric dynamical core using tensor-product compatible Galerkin methods. For the first time, through the combination of a Hamiltonian formulation, compatible Galerkin methods, a specific choice of Galerkin spaces and an energy-conserving time integrator, all of the desirable properties listed in Section 3.1 have been achieved. Several choices for prognostic thermodynamic variable (S versus s), choice of finite element space for s when it is predicted ($\mathbb{W}_0$ or $\mathbb{W}_2$) and representation of the non-linear PV flux term (direct and using q) were explored. In terms of the desirable properties, only the last choice impacts them: use of a direct representation does not allow conservation of potential enstrophy (for the implied shallow water case when $s = g$) and potential vorticity compatibility. The conserved quantities $\mathcal{PV}$, $\mathcal{H}$ and $\mathcal{B}$ have different forms for various choices, but they always exist. The balance, convergence and conservation properties of the scheme were demonstrated by the three test cases in Section 5. For these test cases, there appears to be little difference between the variants introduced here. Of particular interest is the ability of the schemes to conserve nonlinear invariants $\mathcal{H}$ and $\mathcal{B}$ to machine precision, even when using the suboptimal simplified Jacobian in the quasi-Newton solver; and the robustness and stability of the schemes without any added dissipation.

Given the similarities in Hamiltonian structure between the thermal shallow water equations and the fully compressible equations (both hydrostatic and non-hydrostatic variants, with various choices for vertical coordinates), it is anticipated that many of the developments in this paper will immediately carry over to those equations. Preliminary results on a fully compressible quasi-Hamiltonian model in Eulerian coordinates built using these ideas indicate that the key results of fully discrete conservation properties and other desirable characteristics carry over, and will be reported in a future publication. Possible future work on the thermal shallow water equations includes further investigation into optimal simplified Jacobians and preconditioners; the extension of these ideas to domains with boundaries (following the approach in [42]); and the extension to spherical and spheroidal domains with complete representations of the Coriolis force and meridionally varying gravitational potentials. A more detailed study of the Hollingsworth instability for compatible Galerkin methods (especially comparing the two variants of nonlinear PV flux term) is also being undertaken.

7. Acknowledgements

We are grateful to Colin Cotter and Almut Gassmann for their detailed and thorough reviews, which greatly improved the readability and presentation of this paper; and to Colin Cotter, Tom Melvin and Golo Wimmer for many illuminating discussions about the application of Hamiltonian and compatible Galerkin methods for geophysical fluids. We are also thankful to David Ham, Lawrence Mitchell and Miklos Homolya for assistance with development of Themis, without which this work would have been much harder. Christopher Eldred was supported by the French National Research Agency through contract ANR-14-CE23-0010 (HEAT).

8. References

[1] P. Ripa, Conservation laws for primitive equations models with inhomogeneous layers, Geophysical & Astrophysical Fluid Dynamics 70 (1-4) (1993) 85–111. arXiv:https://doi.org/10.1080/03091929308203588, doi:10.1080/03091929308203588.
URL https://doi.org/10.1080/03091929308203588

[2] T. Dubos, S. Dubey, M. Tort, R. Mittal, Y. Meurdesoif, F. Hourdin, Dynamico-1.0, an icosahedral hydrostatic dynamical core designed for consistency and versatility, Geoscientific Model Development 8 (10) (2015) 3131–3150. doi:10.5194/gmd-8-3131-2015.
URL https://www.geosci-model-dev.net/8/3131/2015/

[3] T. Dubos, M. Tort, Equations of atmospheric motion in non-eulerian vertical coordinates: Vector-invariant form and quasi-hamiltonian formulation, Monthly Weather Review 142 (10) (2014) 3860–3880. arXiv:https://doi.org/10.1175/MWR-D-14-00069.1, doi:10.1175/MWR-D-14-00069.1.
URL https://doi.org/10.1175/MWR-D-14-00069.1

[4] P. J. Dellar, Common hamiltonian structure of the shallow water equations with horizontal temperature gradients and magnetic fields, Physics of Fluids 15 (2) (2003) 292–297. arXiv:https://doi.org/10.1063/1.1530576, doi:10.1063/1.1530576.
URL https://doi.org/10.1063/1.1530576

[5] E. S. Warneford, P. J. Dellar, The quasi-geostrophic theory of the thermal shallow water equations, Journal of Fluid Mechanics 723 (2013) 374–403. doi:10.1017/jfm.2013.101.

[6] E. Gouzien, N. Lahaye, V. Zeitlin, T. Dubos, Thermal instability in rotating shallow water with horizontal temperature/density gradients, Physics of Fluids 29 (10) (2017) 101702. arXiv:https://doi.org/10.1063/1.4996981, doi:10.1063/1.4996981.
URL https://doi.org/10.1063/1.4996981

[7] E. S. Warneford, P. J. Dellar, Thermal shallow water models of geostrophic turbulence in jovian atmospheres, Physics of Fluids 26 (1) (2014) 016603. arXiv:https://doi.org/10.1063/1.4861123, doi:10.1063/1.4861123.
URL https://doi.org/10.1063/1.4861123

[8] E. S. Warneford, P. J. Dellar, Super- and sub-rotating equatorial jets in shallow water models of jovian atmospheres: Newtonian cooling versus rayleigh friction, Journal of Fluid Mechanics 822 (2017) 484–511. doi:10.1017/jfm.2017.232.

[9] C. Cotter, J. Shipton, Mixed finite elements for numerical weather prediction, Journal of Computational Physics 231 (21) (2012) 7076 – 7091. doi:https://doi.org/10.1016/j.jcp.2012.05.020.

[10] A. T. T. McRae, C. J. Cotter, Energy- and enstrophy-conserving schemes for the shallow-water equations, based on mimetic finite elements, Quarterly Journal of the Royal Meteorological Society 140 (684) (2014) 2223–2234. doi:10.1002/qj.2291.
URL http://dx.doi.org/10.1002/qj.2291

[11] A. Natale, J. Shipton, C. J. Cotter, Compatible finite element spaces for geophysical fluid dynamics, Dynamics and Statistics of the Climate System 1 (1) (2016) dzw005. doi:10.1093/climsys/dzw005.

[12] A. Arakawa, V. R. Lamb, A potential enstrophy and energy conserving scheme for the shallow water equations, Monthly Weather Review 109 (1) (1981) 18–36. arXiv:https://doi.org/10.1175/1520-0493(1981)109<0018:APEAEC>2.0.CO;2, doi:10.1175/1520-0493(1981)109<0018:APEAEC>2.0.CO;2.
URL https://doi.org/10.1175/1520-0493(1981)109<0018:APEAEC>2.0.CO;2

[13] A. L. Stewart, P. J. Dellar, An energy and potential enstrophy conserving numerical scheme for the multi-layer shallow water equations with complete coriolis force, Journal of Computational Physics 313 (2016) 99 – 120. doi:https://doi.org/10.1016/j.jcp.2015.12.042.
URL http://www.sciencedirect.com/science/article/pii/S002199911500861X

[14] C. Eldred, D. Randall, Total energy and potential enstrophy conserving schemes for the shallow water equations using hamiltonian methods – part 1: Derivation and properties, Geoscientific Model Development 10 (2) (2017) 791–810. doi:10.5194/gmd-10-791-2017.
URL https://www.geosci-model-dev.net/10/791/2017/

[15] T. Ringler, J. Thuburn, J. Klemp, W. Skamarock, A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured c-grids, Journal of Computational Physics 229 (9) (2010) 3065 – 3090. doi:https://doi.org/10.1016/j.jcp.2009.12.007.
URL http://www.sciencedirect.com/science/article/pii/S0021999109006780

[16] R. Sadourny, The dynamics of finite-difference models of the shallow-water equations, Journal of the Atmospheric Sciences 32 (4) (1975) 680–689. arXiv:https://doi.org/10.1175/1520-0469(1975)032<0680:TDOFDM>2.0.CO;2, doi:10.1175/1520-0469(1975)032<0680:TDOFDM>2.0.CO;2.
URL https://doi.org/10.1175/1520-0469(1975)032<0680:TDOFDM>2.0.CO;2

[17] J. Thuburn, C. J. Cotter, A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids, SIAM Journal on Scientific Computing 34 (3) (2012) B203–B225. arXiv:https://doi.org/10.1137/110850293, doi:10.1137/110850293.
URL https://doi.org/10.1137/110850293

[18] M. A. Taylor, A. Fournier, A compatible and conservative spectral element method on unstructured grids, Journal of Computational Physics 229 (17) (2010) 5879 – 5895. doi:https://doi.org/10.1016/j.jcp.2010.04.008.
URL http://www.sciencedirect.com/science/article/pii/S0021999110001841

[19] D. Lee, A. Palha, M. Gerritsma, Discrete conservation properties for shallow water flows using mixed mimetic spectral elements, Journal of Computational Physics 357 (2018) 282 – 304. doi:https://doi.org/10.1016/j.jcp.2017.12.022.
URL http://www.sciencedirect.com/science/article/pii/S0021999117309166

[20] J. Shipton, T. Gibson, C. Cotter, Higher-order compatible finite element schemes for the nonlinear rotating shallow water equations on the sphere, Journal of Computational Physics 375 (2018) 1121 – 1137. doi:https://doi.org/10.1016/j.jcp.2018.08.027.
URL http://www.sciencedirect.com/science/article/pii/S0021999118305503

[21] J. Thuburn, C. J. Cotter, A primal–dual mimetic finite element scheme for the rotating shallow water equations on polygonal spherical meshes, Journal of Computational Physics 290 (Supplement C) (2015) 274 – 297. doi:https://doi.org/10.1016/j.jcp.2015.02.045.
URL http://www.sciencedirect.com/science/article/pii/S0021999115001151

[22] H. Yamazaki, J. Shipton, M. J. Cullen, L. Mitchell, C. J. Cotter, Vertical slice modelling of nonlinear eady waves using a compatible finite element method, Journal of Computational Physics 343 (Supplement C) (2017) 130 – 149. doi:https://doi.org/10.1016/j.jcp.2017.04.006.
URL http://www.sciencedirect.com/science/article/pii/S0021999117302772

[23] P. S. Peixoto, Accuracy analysis of mimetic finite volume operators on geodesic grids and a consistent alternative, Journal of Computational Physics 310 (Supplement C) (2016) 127 – 160. doi:https://doi.org/10.1016/j.jcp.2015.12.058.
URL http://www.sciencedirect.com/science/article/pii/S0021999116000267

[24] C. Eldred, D. Y. L. Roux, Dispersion analysis of compatible galerkin schemes for the 1d shallow water model, Journal of Computational Physics 371 (2018) 779 – 800. doi:https://doi.org/10.1016/j.jcp.2018.06.007.
URL http://www.sciencedirect.com/science/article/pii/S0021999118303851

[25] A. Gassmann, Inspection of hexagonal and triangular c-grid discretizations of the shallow water equations, Journal of Computational Physics 230 (7) (2011) 2706 – 2721.

doi:https://doi.org/10.1016/j.jcp.2011.01.014.
URL http://www.sciencedirect.com/science/article/pii/S0021999111000325

[26] T. Melvin, Dispersion analysis of the pn-pn-1dg mixed finite element pair for atmospheric modelling, Journal of Computational Physics 355 (Supplement C) (2018) 342 – 365. doi:https://doi.org/10.1016/j.jcp.2017.11.019.
URL http://www.sciencedirect.com/science/article/pii/S0021999117308574

[27] T. Melvin, A. Staniforth, C. Cotter, A two-dimensional mixed finite-element pair on rectangles, Quarterly Journal of the Royal Meteorological Society 140 (680) (2014) 930–942. doi:10.1002/qj.2189.

[28] T. Melvin, J. Thuburn, Wave dispersion properties of compound finite elements, Journal of Computational Physics 338 (2017) 68 – 90. doi:https://doi.org/10.1016/j.jcp.2017.02.025.
URL http://www.sciencedirect.com/science/article/pii/S0021999117301183

[29] V. Rostand, D. Y. L. Roux, G. Carey, Kernel analysis of the discretized finite difference and finite element shallow-water models, SIAM Journal on Scientific Computing 31 (1) (2008) 531–556. doi:10.1137/070695198.

[30] D. Y. L. Roux, Spurious inertial oscillations in shallow-water models, Journal of Computational Physics 231 (24) (2012) 7959 – 7987. doi:https://doi.org/10.1016/j.jcp.2012.04.052.

[31] A. Staniforth, T. Melvin, C. Cotter, Analysis of a mixed finite-element pair proposed for an atmospheric dynamical core, Quarterly Journal of the Royal Meteorological Society 139 (674) (2013) 1239–1254. doi:10.1002/qj.2028.

[32] J. Thuburn, Numerical wave propagation on the hexagonal c-grid, Journal of Computational Physics 227 (11) (2008) 5836 – 5858. doi:https://doi.org/10.1016/j.jcp.2008.02.010.
URL http://www.sciencedirect.com/science/article/pii/S0021999108001162

[33] S. Ničković, M. B. Gavrilov, I. A. Tošić, Geostrophic adjustment on hexagonal grids, Monthly Weather Review 130 (3) (2002) 668–683. arXiv:https://doi.org/10.1175/1520-0493(2002)130<0668:GAOHG>2.0.CO;2, doi:10.1175/1520-0493(2002)130<0668:GAOHG>2.0.CO;2.
URL https://doi.org/10.1175/1520-0493(2002)130<0668:GAOHG>2.0.CO;2

[34] J. Thuburn, T. Ringler, W. Skamarock, J. Klemp, Numerical representation of geostrophic modes on arbitrarily structured c-grids, Journal of Computational Physics 228 (22) (2009) 8321 – 8335. doi:https://doi.org/10.1016/j.jcp.2009.08.006.
URL http://www.sciencedirect.com/science/article/pii/S0021999109004434

[35] J. Thuburn, C. J. Cotter, T. Dubos, A mimetic, semi-implicit, forward-in-time, finite volume shallow water model: comparison of hexagonal-icosahedral and cubed-sphere grids, Geoscientific Model Development 7 (3) (2014) 909–929. doi:10.5194/gmd-7-909-2014.
URL https://www.geosci-model-dev.net/7/909/2014/

[36] H. Weller, Non-orthogonal version of the arbitrary polygonal c-grid and a new diamond grid, Geoscientific Model Development 7 (3) (2014) 779–797. doi:10.5194/gmd-7-779-2014.
URL https://www.geosci-model-dev.net/7/779/2014/

[37] M. J. Bell, P. S. Peixoto, J. Thuburn, Numerical instabilities of vector-invariant momentum equations on rectangular c-grids, Quarterly Journal of the Royal Meteorological Society 143 (702) (2017) 563–581. doi:10.1002/qj.2950.
URL http://dx.doi.org/10.1002/qj.2950

[38] A. Hollingsworth, P. Kållberg, V. Renner, D. M. Burridge, An internal symmetric computational instability, Quarterly Journal of the Royal Meteorological Society 109 (460) (1983) 417–428. doi:10.1002/qj.49710946012.
URL http://dx.doi.org/10.1002/qj.49710946012

[39] L. Lazić, Z. Janjić, F. Mesinger, "non-cancellation" instability in horizontal advection schemes for momentum equations, Meteorology and Atmospheric Physics 35 (1) (1986) 49–52. doi:10.1007/BF01029522.
URL https://doi.org/10.1007/BF01029522

[40] P. Peixoto, J. Thuburn, M. Bell, Numerical instabilities of spherical shallow water models considering small equivalent depths, Quarterly Journal of the Royal Meteorological SocietyQJ-17-0211.R1. doi:10.1002/qj.3191.
URL http://dx.doi.org/10.1002/qj.3191

[41] A. Gassmann, Discretization of generalized coriolis and friction terms on the deformed hexagonal c-grid, Quarterly Journal of the Royal Meteorological Society 0 (0). arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3294, doi:10.1002/qj.3294.
URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3294

[42] W. Bauer, C. Cotter, Energy-enstrophy conserving compatible finite element schemes for the rotating shallow water equations with slip boundary conditions, Journal of Computational Physics (2018) –doi:https://doi.org/10.1016/j.jcp.2018.06.071.
URL https://www.sciencedirect.com/science/article/pii/S0021999118304509

[43] R. Salmon, Poisson-bracket approach to the construction of energy- and potential-enstrophy-conserving algorithms for the shallow-water equations, Journal of the Atmospheric Sciences 61 (16) (2004) 2016–2036.

arXiv:https://doi.org/10.1175/1520-0469(2004)061<2016:PATTCO>2.0.CO;2, doi:10.1175/1520-0469(2004)061<2016:PATTCO>2.0.CO;2.
URL https://doi.org/10.1175/1520-0469(2004)061<2016:PATTCO>2.0.CO;2

[44] M. Sommer, P. Névir, A conservative scheme for the shallow-water system on a staggered geodesic grid based on a nambu representation, Quarterly Journal of the Royal Meteorological Society 135 (639) (2009) 485–494. doi:10.1002/qj.368.
URL http://dx.doi.org/10.1002/qj.368

[45] A. Gassmann, A global hexagonal c-grid non-hydrostatic dynamical core (icon-iap) designed for energetic consistency, Quarterly Journal of the Royal Meteorological Society 139 (670) (2013) 152–175. doi:10.1002/qj.1960.
URL http://dx.doi.org/10.1002/qj.1960

[46] A. Gassmann, H.-J. Herzog, Towards a consistent numerical compressible non-hydrostatic model using generalized hamiltonian tools, Quarterly Journal of the Royal Meteorological Society 134 (635) (2008) 1597–1613. doi:10.1002/qj.297.
URL http://dx.doi.org/10.1002/qj.297

[47] M. Tort, T. Dubos, T. Melvin, Energy-conserving finite-difference schemes for quasi-hydrostatic equations, Quarterly Journal of the Royal Meteorological Society 141 (693) (2015) 3056–3075. doi:10.1002/qj.2590.
URL http://dx.doi.org/10.1002/qj.2590

[48] M. D. Toy, R. D. Nair, A potential enstrophy and energy conserving scheme for the shallow-water equations extended to generalized curvilinear coordinates, Monthly Weather Review 145 (3) (2017) 751–772. arXiv:https://doi.org/10.1175/MWR-D-16-0250.1, doi:10.1175/MWR-D-16-0250.1.
URL https://doi.org/10.1175/MWR-D-16-0250.1

[49] E. Kritsikis, T. Dubos, Higher-order finite elements for the shallow-water equations on the cubed sphere (2014).

[50] T. G. Shepherd, A unified theory of available potential energy, Atmosphere-Ocean 31 (1) (1993) 1–26. arXiv:https://doi.org/10.1080/07055900.1993.9649460, doi:10.1080/07055900.1993.9649460.
URL https://doi.org/10.1080/07055900.1993.9649460

[51] P. Ripa, Linear waves in a one-layer ocean model with thermodynamics, Journal of Geophysical Research: Oceans 101 (C1) (1996) 1233–1245. doi:10.1029/95JC02899.
URL http://dx.doi.org/10.1029/95JC02899

[52] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, Acta Numerica 15 (2006) 1–155. doi:10.1017/S0962492906210018.

[53] R. Hiemstra, D. Toshniwal, R. Huijsmans, M. Gerritsma, High order geometric methods with exact conservation properties, Journal of Computational Physics 257 (Part B) (2014) 1444 – 1471, physics-compatible numerical methods. doi:https://doi.org/10.1016/j.jcp.2013.09.027.

[54] A. Staniforth, J. Thuburn, Horizontal grids for global weather and climate prediction models: a review, Quarterly Journal of the Royal Meteorological Society 138 (662) (2012) 1–26. doi:10.1002/qj.958.
URL http://dx.doi.org/10.1002/qj.958

[55] D. Arnold, R. Falk, R. Winther, Finite element exterior calculus: from hodge theory to numerical stability, Bulletin of the American mathematical society 47 (2) (2010) 281–354.

[56] S. Danilov, On utility of triangular c-grid type discretization for numerical modeling of large-scale ocean flows, Ocean Dynamics 60 (6) (2010) 1361–1369. doi:10.1007/s10236-010-0339-6.
URL https://doi.org/10.1007/s10236-010-0339-6

[57] A. McRae, G. Bercea, L. Mitchell, D. Ham, C. Cotter, Automated generation and symbolic manipulation of tensor product finite elements, SIAM Journal on Scientific Computing 38 (5) (2016) S25–S47. arXiv:https://doi.org/10.1137/15M1021167, doi:10.1137/15M1021167.
URL https://doi.org/10.1137/15M1021167

[58] A. Buffa, J. Rivas, G. Sangalli, R. Vazquez, Isogeometric discrete differential forms in three dimensions, SIAM Journal on Numerical Analysis 49 (2) (2011) 818–844. doi:10.1137/100786708.

[59] A. Buffa, G. Sangalli, R. Vazquez, Isogeometric methods for computational electromagnetics: B-spline and t-spline discretizations, Journal of Computational Physics 257 (Part B) (2014) 1291 – 1320, physics-compatible numerical methods. doi:https://doi.org/10.1016/j.jcp.2013.08.015.

[60] J. Banks, T. Hagstrom, On galerkin difference methods, Journal of Computational Physics 313 (Supplement C) (2016) 310 – 327. doi:https://doi.org/10.1016/j.jcp.2016.02.042.

[61] A. Sarmiento, A. Cortes, D. Garcia, L. Dalcin, N. Collier, V. Calo, Petiga-mf: A multi-field high-performance toolbox for structure-preserving b-splines spaces, Journal of Computational Science 18 (Supplement C) (2017) 117 – 131. doi:https://doi.org/10.1016/j.jocs.2016.09.010.

[62] W. Bauer, A new hierarchically-structured n-dimensional covariant form of rotating equations of geophysical fluid dynamics, GEM - International Journal on Geomathe-

matics 7 (1) (2016) 31–101. doi:10.1007/s13137-015-0074-8.
URL https://doi.org/10.1007/s13137-015-0074-8

[63] J. B. Perot, C. J. Zusi, Differential forms for scientists and engineers, Journal of Computational Physics 257 (Part B) (2014) 1373 – 1393, physics-compatible numerical methods. doi:https://doi.org/10.1016/j.jcp.2013.08.007.
URL http://www.sciencedirect.com/science/article/pii/S0021999113005354

[64] E. Tonti, Why starting from differential equations for computational physics?, Journal of Computational Physics 257 (Part B) (2014) 1260 – 1290, physics-compatible numerical methods. doi:https://doi.org/10.1016/j.jcp.2013.08.016.
URL http://www.sciencedirect.com/science/article/pii/S0021999113005548

[65] M. E. Rognes, D. A. Ham, C. J. Cotter, A. T. T. McRae, Automating the solution of pdes on the sphere and other manifolds in fenics 1.2, Geoscientific Model Development 6 (6) (2013) 2099–2119. doi:10.5194/gmd-6-2099-2013.
URL https://www.geosci-model-dev.net/6/2099/2013/

[66] H. Heumann, R. Hiptmair, C. Pagliantini, Stabilized galerkin for transient advection of differential forms, Discrete and Continuous Dynamical Systems 9 (2016) 185. doi:10.3934/dcdss.2016.9.185.

[67] A. Natale, C. J. Cotter, Scale-selective dissipation in energy-conserving finite-element schemes for two-dimensional turbulence, Quarterly Journal of the Royal Meteorological Society 143 (705) (2017) 1734–1745. doi:10.1002/qj.3063.
URL http://dx.doi.org/10.1002/qj.3063

[68] A. Natale, C. J. Cotter, A variational $\boldsymbol{H}(\mathrm{div})$ finite-element discretization approach for perfect incompressible fluids, IMA Journal of Numerical Analysis (2017) drx033doi:10.1093/imanum/drx033.

[69] R. I. McLachlan, Spatial discretization of partial differential equations with integrals, IMA Journal of Numerical Analysis 23 (4) (2003) 645–664. doi:10.1093/imanum/23.4.645.

[70] D. Cohen, E. Hairer, Linear energy-preserving integrators for poisson systems, BIT Numerical Mathematics 51 (1) (2011) 91–101. doi:10.1007/s10543-011-0310-z.
URL https://doi.org/10.1007/s10543-011-0310-z

[71] T. Benacchio, N. Wood, Semi-implicit semi-lagrangian modelling of the atmosphere: A met office perspective, Communications in Applied and Industrial Mathematics 7 (2016) 4–25.

[72] N. Wood, A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat, J. Thuburn, An inherently mass-conserving semi-implicit semi-lagrangian discretization of the deep-atmosphere global non-hydrostatic

equations, Quarterly Journal of the Royal Meteorological Society 140 (682) (2014) 1505–1520. doi:10.1002/qj.2235.
URL http://dx.doi.org/10.1002/qj.2235

[73] C. Eldred, Themis web page (2018).
URL https://github.com/celdred/themis

[74] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. Mcrae, G.-T. Bercea, G. R. Markall, P. H. J. Kelly, Firedrake: Automating the finite element method by composing abstractions, ACM Trans. Math. Softw. 43 (3) (2016) 24:1–24:27. doi:10.1145/2998441.
URL http://doi.acm.org/10.1145/2998441

[75] M. S. Alnaes, A. Logg, K. B. Olgaard, M. E. Rognes, G. N. Wells, Unified form language: A domain-specific language for weak formulations of partial differential equations, ACM Trans. Math. Softw. 40 (2) (2014) 9:1–9:37. doi:10.1145/2566630.
URL http://doi.acm.org/10.1145/2566630

[76] M. Homolya, R. C. Kirby, D. A. Ham, Exposing and exploiting structure: optimal code generation for high-order finite element methods, CoRR abs/1711.02473. arXiv:1711.02473.
URL http://arxiv.org/abs/1711.02473

[77] M. Homolya, L. Mitchell, F. Luporini, D. A. Ham, TSFC: a structure-preserving form compiler, CoRR abs/1705.03667. arXiv:1705.03667.
URL http://arxiv.org/abs/1705.03667

[78] Finat web page (2018).
URL https://github.com/FInAT/FInAT

[79] F. Luporini, D. A. Ham, P. H. J. Kelly, An algorithm for the optimization of finite element integration loops, ACM Trans. Math. Softw. 44 (1) (2017) 3:1–3:26. doi:10.1145/3054944.
URL http://doi.acm.org/10.1145/3054944

[80] F. Luporini, A. L. Varbanescu, F. Rathgeber, G.-T. Bercea, J. Ramanujam, D. A. Ham, P. H. J. Kelly, Cross-loop optimization of arithmetic intensity for finite element local assembly, ACM Trans. Archit. Code Optim. 11 (4) (2015) 57:1–57:25. doi:10.1145/2687415.
URL http://doi.acm.org/10.1145/2687415

[81] R. C. Kirby, Algorithm 839: Fiat, a new paradigm for computing finite element basis functions, ACM Trans. Math. Softw. 30 (4) (2004) 502–516. doi:10.1145/1039813.1039820.
URL http://doi.acm.org/10.1145/1039813.1039820

[82] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, R. T. Mills, T. Munson, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, PETSc Web page (2018).
URL http://www.mcs.anl.gov/petsc

[83] D. L. Williamson, J. B. Drake, J. J. Hack, R. Jakob, P. N. Swarztrauber, A standard test set for numerical approximations to the shallow water equations in spherical geometry, Journal of Computational Physics 102 (1) (1992) 211 – 224. doi:https://doi.org/10.1016/S0021-9991(05)80016-6.
URL http://www.sciencedirect.com/science/article/pii/S0021999105800166

[84] M. Giorgetta, T. Hundertmark, P. Korn, S. Reich, M. Restelli, Conservative Space and Time Regularizations for the ICON Model, Reports on Earth System Science.
URL http://publications.imp.fu-berlin.de/770/

[85] T. Dubos, A variational formulation of geophysical fluid motion in non-eulerian coordinates, Quarterly Journal of the Royal Meteorological Society 143 (702) (2017) 542–551. doi:10.1002/qj.2942.
URL http://dx.doi.org/10.1002/qj.2942

Appendix A. Thermogeostrophic Balance

Start with the equation of thermogeostrophic balance [6] (with $b = 0$)

$$f\,\mathbf{u}^T + s\nabla h + \frac{h}{2}\nabla s = 0 \tag{A.1}$$

Now consider the case where $\mathbf{u}$ and h satisfy geostrophic balance

$$f\,\mathbf{u}^T + g\nabla h = 0 \tag{A.2}$$

Then thermogeostrophic balance can be rewritten as

$$2\frac{(s-g)}{h}\nabla h + \nabla s = 0 \tag{A.3}$$

Given h, this can be solved for s. Additionally, it is clear that if $s = g$, this will be identically zero, and thermogeostrophic balance will reduce to geostrophic balance. Unlike the shallow water equations and geostrophic balance, it is not clear if the discretization presented in this paper has a discrete analogue of thermogeostrophic balance.

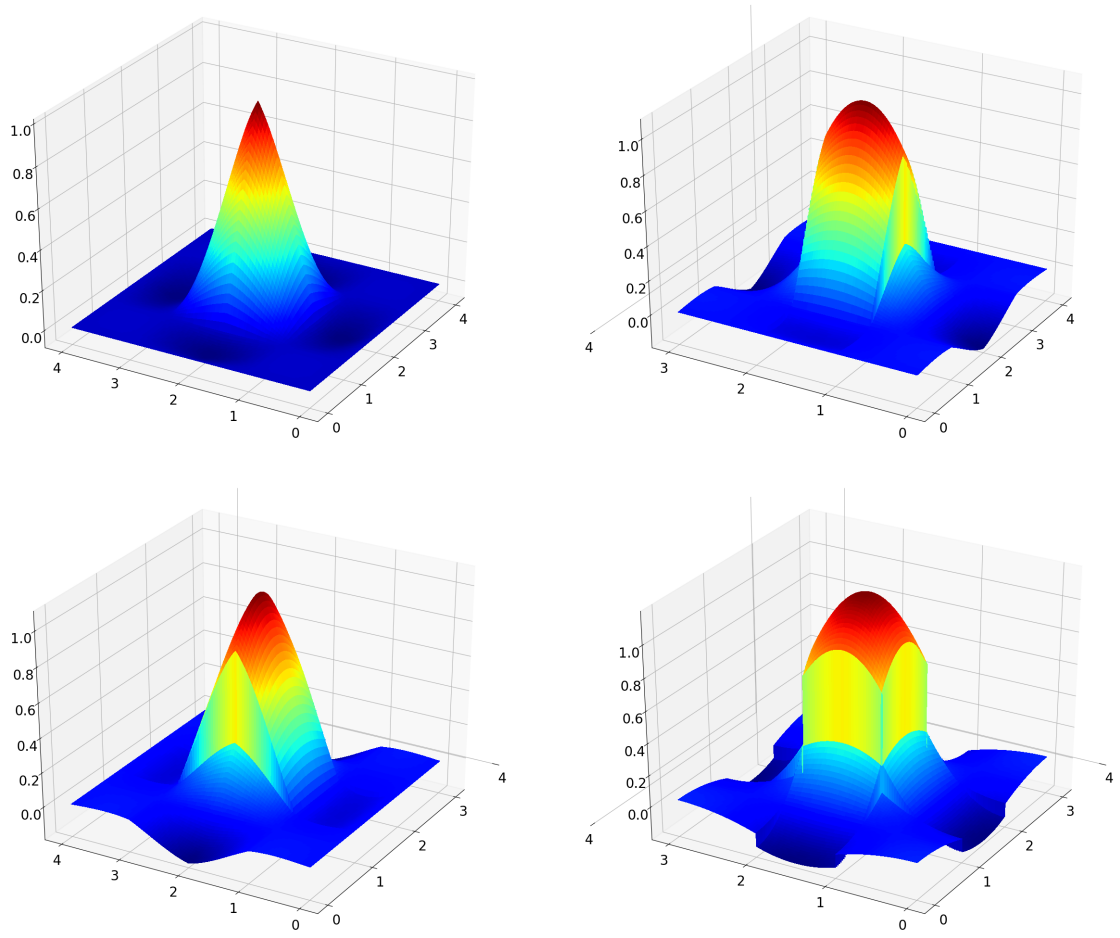

Figure 1: Basis functions for the $\mathbb{W}_0$ (upper left), $\mathbb{W}_1$ (x/u component in upper right, y/v component in lower left) and $\mathbb{W}_2$ (lower right) spaces of the MGD_3 family. Here the element width and basis functions have been normalized to unity. Note that unlike standard mixed finite elements, all of the basis functions in a given space are identical; and the basis functions are not localized to an element or a pair of neighboring elements, but instead have extended support (although they remain local).

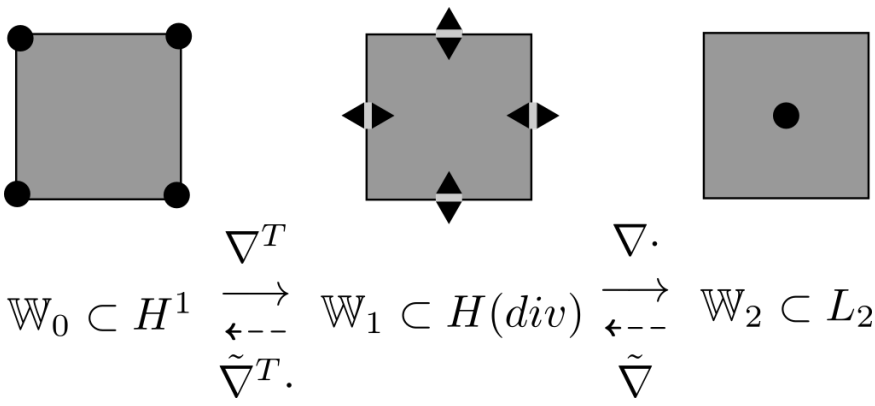

Figure 2: Compatible Galerkin spaces $\mathbb{W}_0$, $\mathbb{W}_1$ and $\mathbb{W}_2$ and corresponding discrete deRham complex in 2D for the MGD_n family. Solid lines indicate strong operators (∇^T, $\nabla\cdot$), while dashed lines indicate weak operators ($\tilde{\nabla}$, $\tilde{\nabla}^T\cdot$). The degrees of freedom are illustrated for the MGD_n family, which are also correct for the lowest order $Q_r^-\Lambda^k$/mimetic spectral element and odd-order isogeometric analysis families.

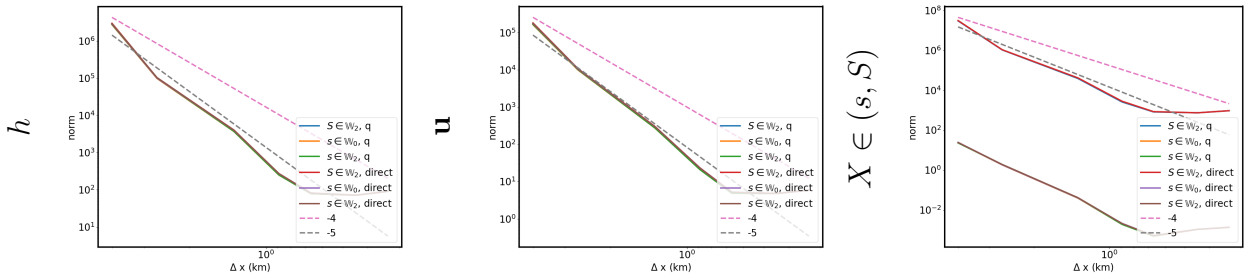

Figure 3: Convergence plot for the zonal geostrophic balance test case, showing the L_2 norms of the difference between the initial $(h, \mathbf{u}, X)$ and the final $(h, \mathbf{u}, X)$ for all 6 variants, where $X \in (s, S)$ is the predicted buoyancy variable. This explains the difference in scales between the errors in the rightmost plot. All of the variants are converging at between 4th and 5th order in $(h, \mathbf{u}, X)$, significantly better than expected.

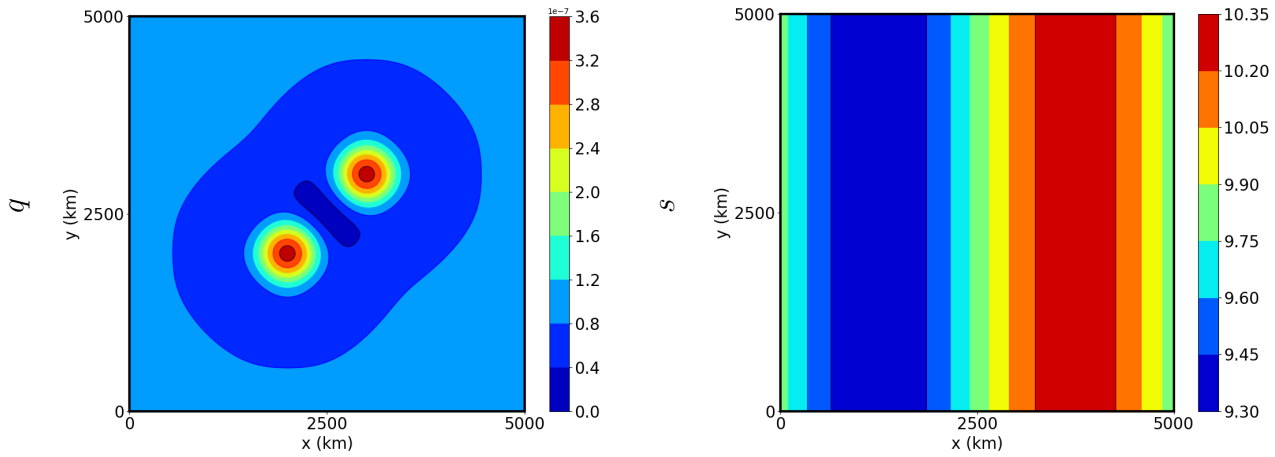

Figure 4: The initial values for potential vorticity q (left) and buoyancy s (right) in the double vortex test case.

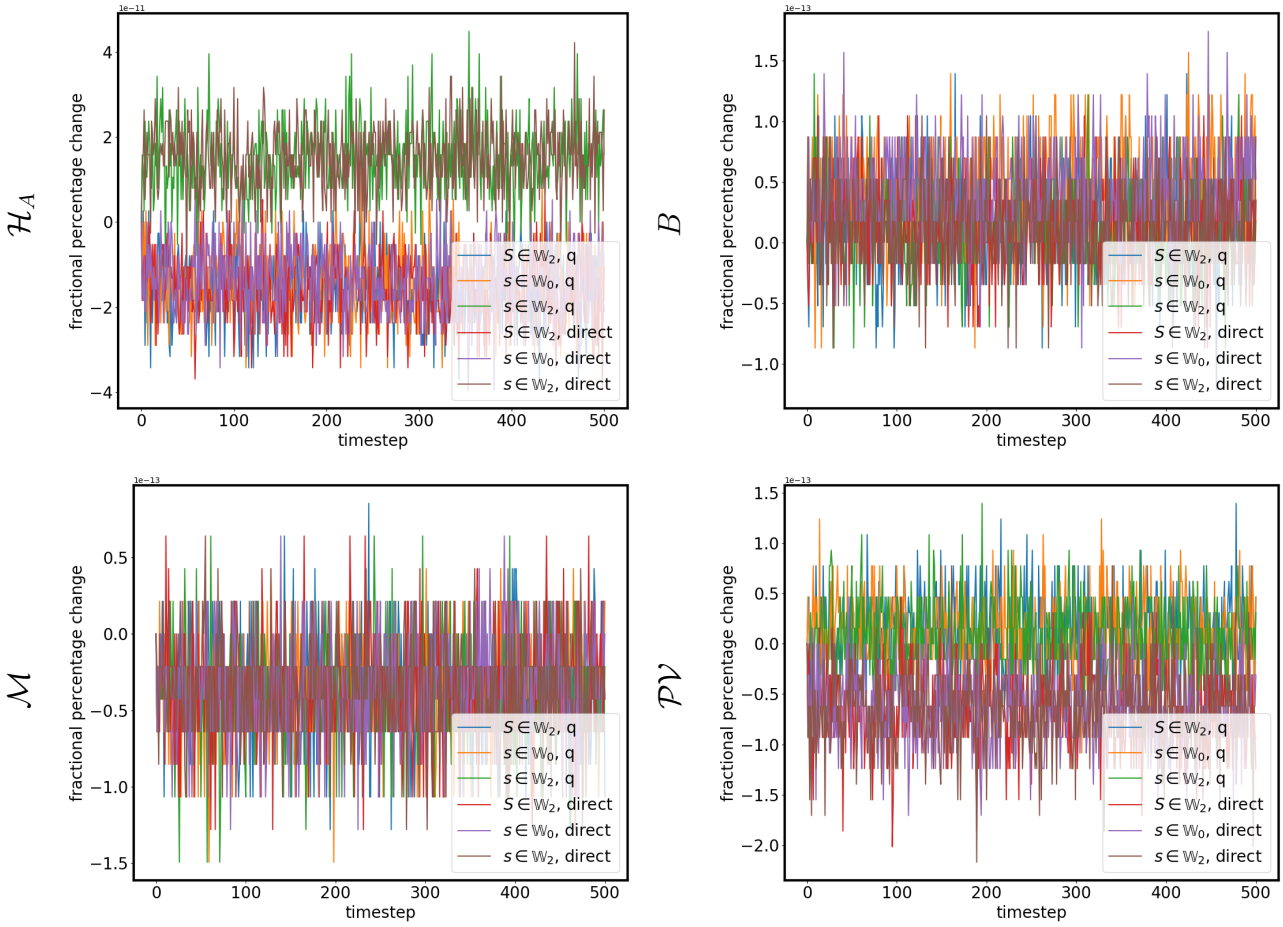

Figure 5: Convergence properties for $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ in the double vortex test case, shown for all six variants. The fractional change ($\frac{x-x_0}{x_0} * 100$) in the relevant quantity versus the time step is plotted. All variants are conserving $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ to machine precision.

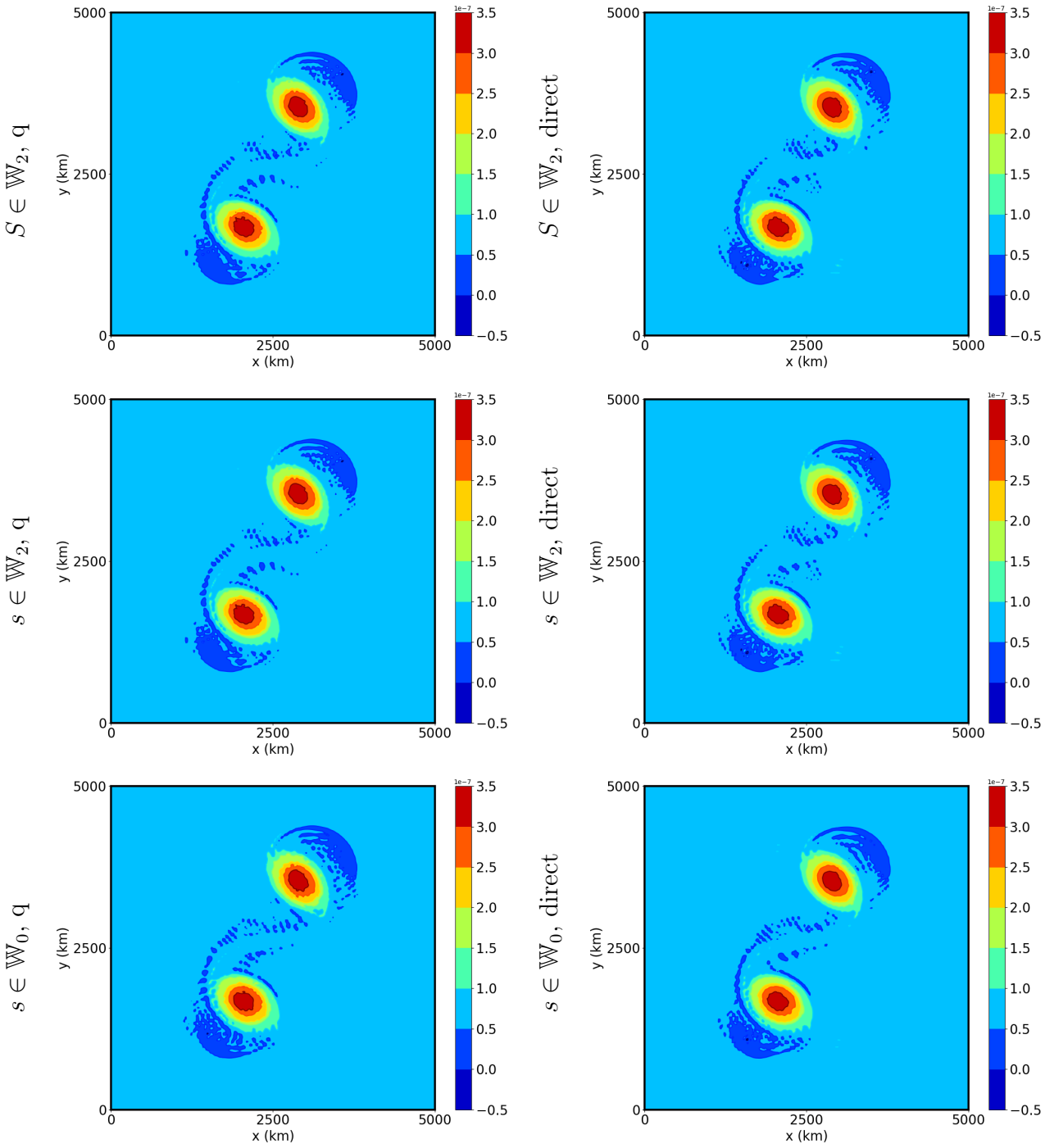

Figure 6: The potential vorticity q at $N = 250$ for the double vortex test case for all six variants. Small scale features have developed following the path of the vortices, although the main vortex cores remain coherent. There is little difference between the six variants.

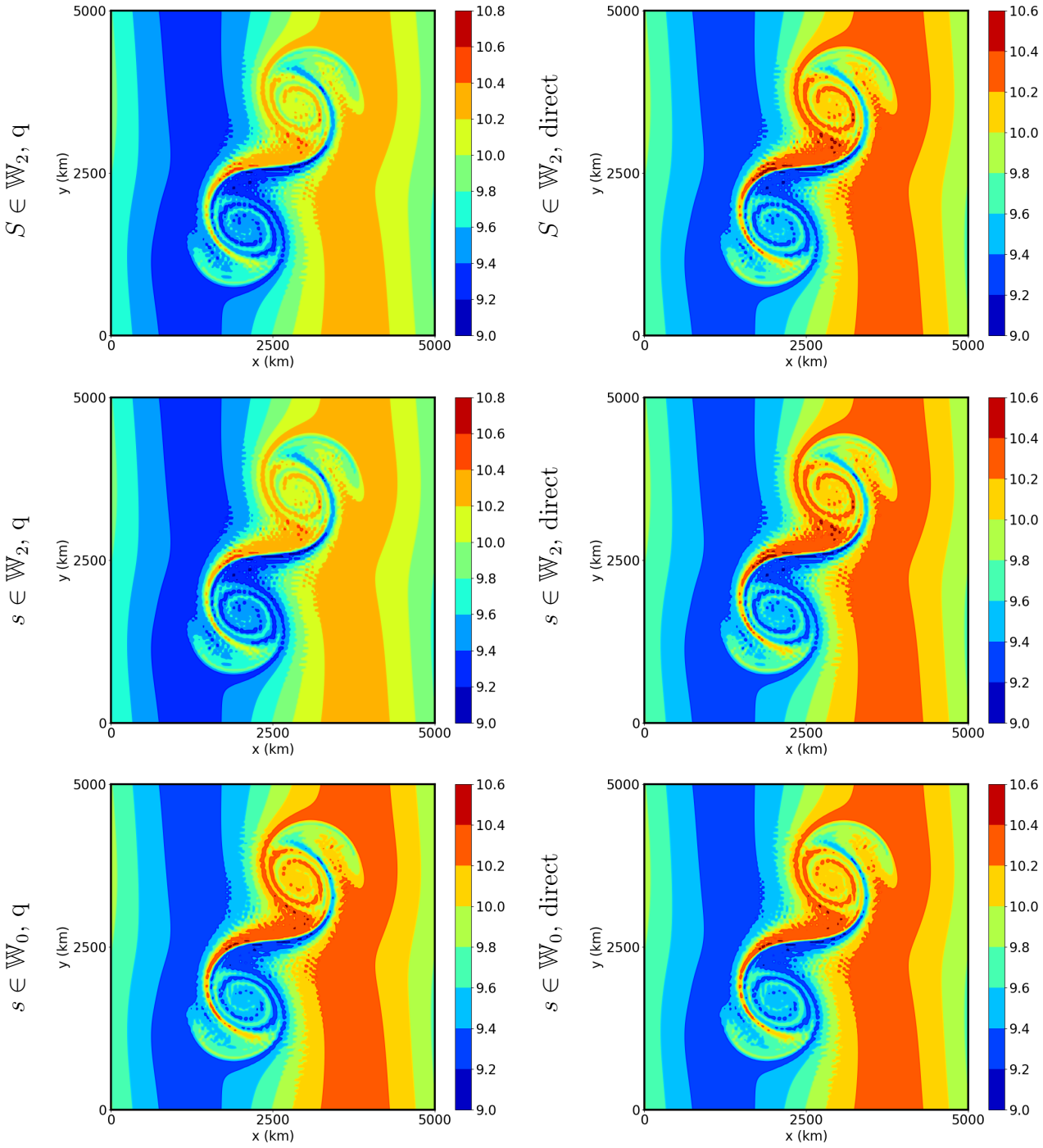

Figure 7: The buoyancy s at $N = 250$ for the double vortex test case for all six variants. The perturbations from the initial s_0 follow the path of the vortices, and small scale features have appeared. As in Figure 6, there is little difference between the six variants.

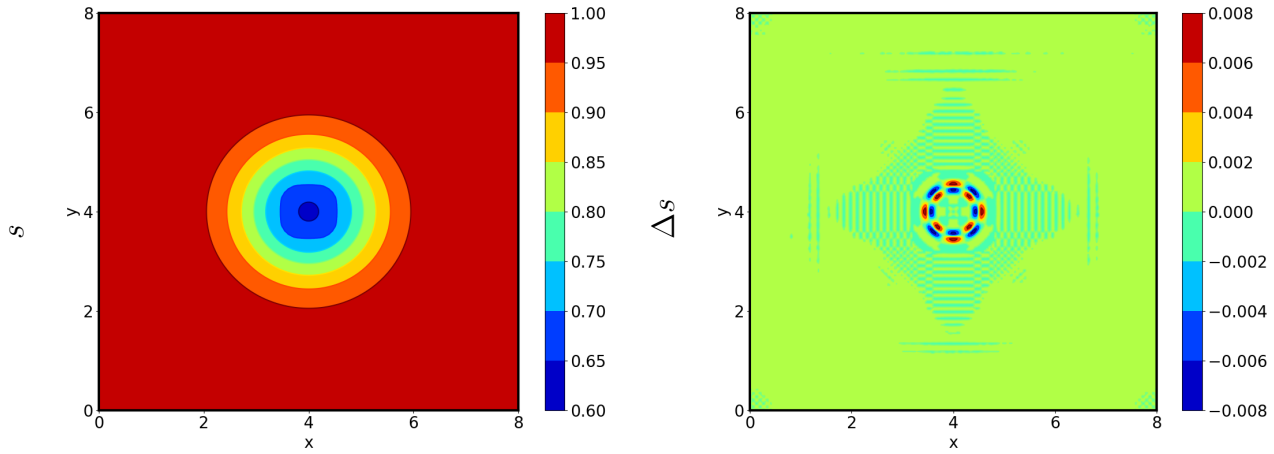

Figure 8: The initial (left) and perturbation (right) values for the buoyancy s in the thermal instability test case. Recall that $l = 4$, and the perburation is localized to a small ring of the domain.

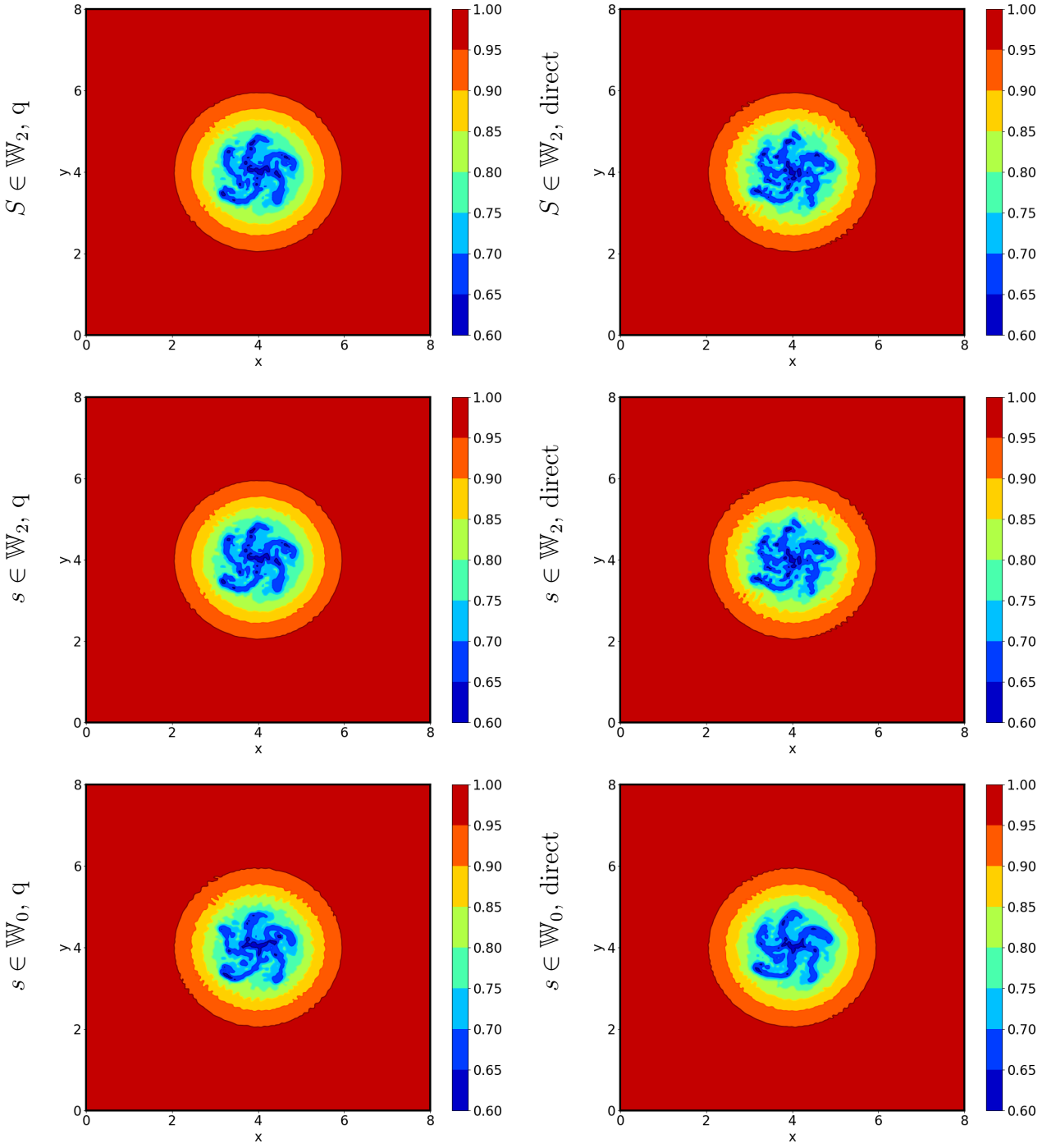

Figure 9: The buoyancy s at $N = 400$ for the thermal instability test case. The wavenumber 4 structure is clearly apparent, and small scale features have developed. Further simulations leads to a complete breakdown of the initial buoyancy and nonlinear saturation of the instability (not shown). It is remarkable that these runs are stable without any added dissipation, even after nonlinear saturation, despite the small scale features. However, it is not unexpected, given the total energy, potential vorticity and buoyancy conserving capabilities of the schemes.

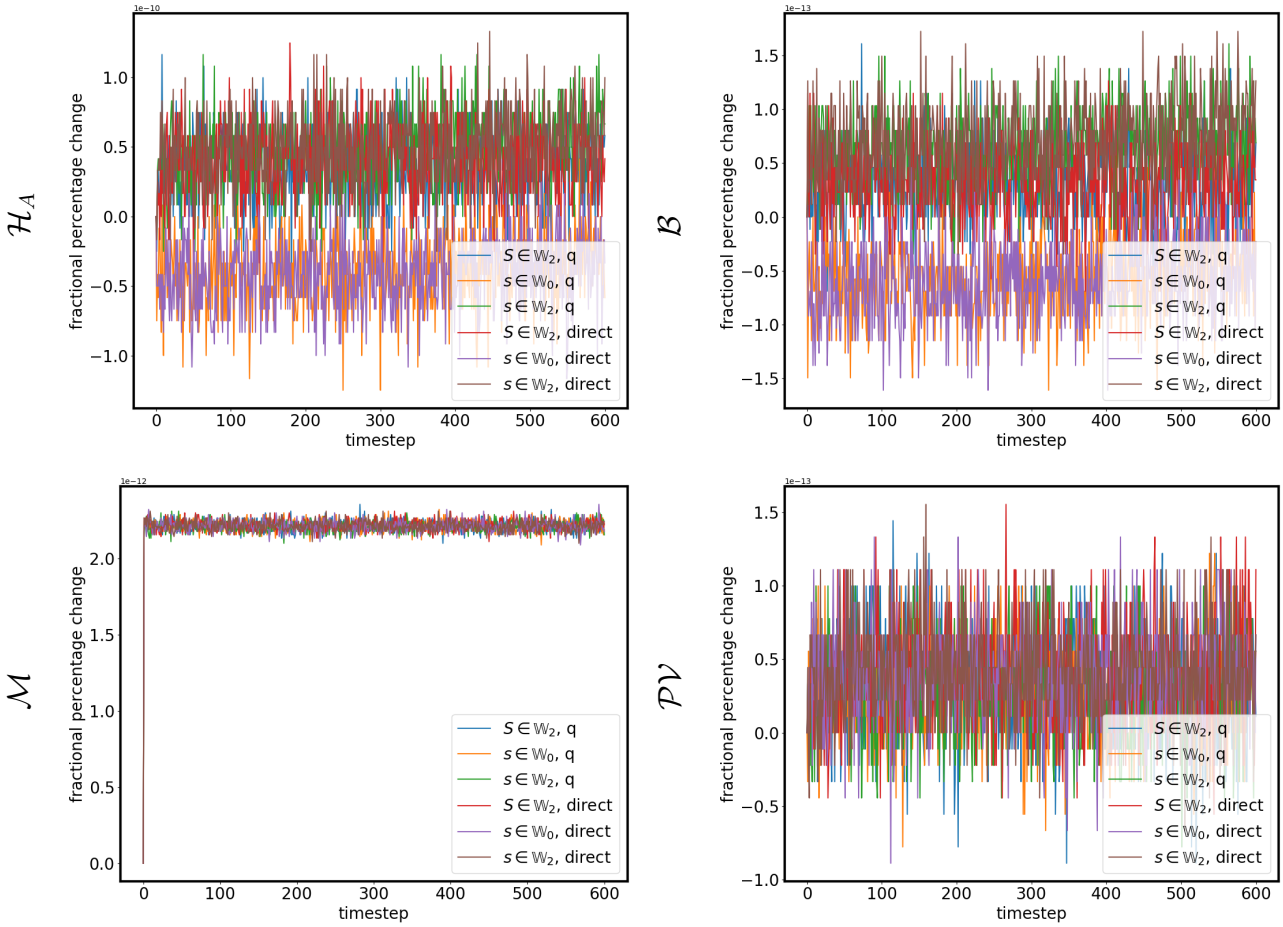

Figure 10: Convergence properties for $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ in the thermal instability test case, shown for all six variants. The fractional change ($\frac{x-x_0}{x_0} * 100$) in the relevant quantity versus the time step is plotted. As in Figure 5, all variants are conserving $\mathcal{M}$, $\mathcal{B}$, $\mathcal{H}_A$ and $\mathcal{PV}$ to machine precision, albeit with a strange (and still unexplained) jump in mass in the first time step for all variants.